

CHAPTER 1

BACKGROUND ON THE EPIDEMIOLOGY AND MODELING OF HIV/AIDS

Before modeling any disease it is crucial to understand the epidemiological features of the disease. The first four sections of this chapter present epidemiological information on the human immunodeficiency virus (HIV) and the acquired immune deficiency syndrome (AIDS) including transmission mechanisms, data, and historical, biological and clinical aspects. The last three sections discuss not only advantages and disadvantages of various approaches to modeling HIV/AIDS, but also the purposes and limitations of epidemiological modeling.

1.1 HIV Transmission Mechanisms

The three known modes of transmission of HIV are sexual contact, direct contact with HIV-infected blood or fluids and perinatal transmission from an infected mother to her child. There is no evidence of spread from toilet seats, by insects or by casual contact. There are indications that the transmissibility of HIV infection varies greatly during the multi-year course of infection in an individual (Goedert, Eyster et al., 1987; Longini et al., 1990). Transmission may be more likely during the early flu-like illness, which occurs just after HIV infection and before the body develops an antibody response. Since the HIV virus level in the blood decreases during the asymptomatic period, the transmissibility appears to be lower in this period. Then the transmissibility seems to increase again as the CD4⁺ cell count gets low, the HIV virus level in the blood increases, and symptoms appear.

In homosexual partners the most important route of transmission seems to be receptive anal intercourse. The likelihood of transmission seems to be less for insertive anal intercourse and for insertive or receptive orogenital contact (Kaslow and Francis, 1989, p. 98). In heterosexual partners engaging in penile-vaginal intercourse, both male-to-female and female-to-male transmissions do occur. However, a recent study in the U.S. suggests that male-to-female HIV transmission is much more likely than female-to-male transmission (Padian et al., 1991). Although HIV has been isolated in saliva, there is strong evidence against HIV transmission by kissing (Kaslow and Francis, 1989, p. 99).

The transfer of blood from an HIV-infected person to another person has been clearly established as a mode of transmission. In the U.S. this has been an important transmission mechanism for the needle-sharing IVDUs. In Africa nonsterile needles used for injection of medications have resulted in numerous HIV infections. Before blood screening began in the U.S. in 1985, some HIV infections resulted from blood transfusions with HIV-infected blood and from HIV-infected blood factor concentrate for hemophiliacs. Other sources of HIV infection in this category are transplanted organs such as kidneys and semen for artificial insemination. Transmission of HIV by arthropod vectors such as mosquitoes has not been observed and seems very unlikely.

Perinatal transmission from an HIV-infected mother to her child at a time before, during, or after birth is the third transmission mode. The exact times of transmission and the role of

breast feeding have not been established, but approximately 30% of the children of HIV-infected mothers are infected.

1.2 Biological and Clinical Aspects of HIV Infection

The retrovirus called *human immunodeficiency virus* (HIV) was established in 1983 as the causative agent of AIDS. Isolates of HIV are molecularly and biologically heterogeneous, with some isolates being more virulent. Moreover, the HIV virus can change rapidly, even within an individual. The HIV infects a subpopulation of thymus-derived T lymphocytes called CD4⁺ lymphocytes or T4 cells, which are helper/inducer cells. These T cells perform recognition and induction functions as part of the immune response to foreign stimuli. The HIV integrates into the CD4⁺ host cell DNA, where it can remain dormant for a long time. The CD4⁺ T lymphocytes are eventually killed by the HIV while the HIV reproduces itself, so that the number of CD4⁺ cells gradually decreases from the normal number of about 900/ml (Kaslow and Francis, 1989). This leads to severe immunodeficiency in persons infected with HIV. Thus the natural history of HIV infection is a gradual depletion of CD4⁺ cells, progressive unresponsiveness of the immune system and increased susceptibility to opportunistic infections such as *Pneumocystis Carinii* pneumonia and malignancies such as Kaposi's sarcoma.

Transmission of HIV infection can be through the transfer of either cell-free HIV virus or HIV-infected lymphocytes. The probability of transmission appears to depend on the stage of HIV infection. In the first few weeks following infection, more HIV has been isolated from blood plasma than in the asymptomatic stage, so that people in the pre-antibody stage may be more infectious than people in the asymptomatic stage. In the late stages of HIV infection and AIDS, the cell-free HIV virus is found more frequently in blood plasma, so that these people may also be more infectious.

The enzyme-linked immunoassay (ELISA) test for HIV antibodies is inexpensive and useful for screening large numbers of samples; this test has almost no false negatives, but it is very sensitive (i.e., it has false positives at the rate of about 2 per 1000). Samples which are positive by the ELISA test are then tested by the Western blot assay. This immunoblot method is very specific (i.e., false positives occur in no more than 0.001% of the samples). Another test used is an indirect immunofluorescent antibody test involving microscopic examination of infected cell spots on glass slides. The combination of these and other tests such as PCR (polymerase chain reaction) are quite accurate in identifying HIV-positive individuals (Kaslow and Francis, 1989).

There are many different clinical manifestations of HIV infection and AIDS. Acute HIV infection occurs just after infection in some patients and is characterized as an acute febrile illness with fever, sweats, lethargy, muscle ache, headache and sore throat. These symptoms may last 2 to 3 weeks. About 2 months after infection, the immune system has generated antibodies which are recognized by the ELISA test. After an asymptomatic period of about 5 years, an HIV-infected person may develop some symptoms such as oral candidiasis (thrush), hairy leukoplakia, herpes zoster, weight loss, diarrhea, persistent generalized lymphadenopathy, neurologic diseases (dementia, myelopathy, polyneuropathy) and tuberculosis. The HIV-infected person may die of these diseases, but most get one or more of the opportunistic infections or

neoplasms which characterize AIDS. These include cytomegalovirus infection, *Pneumocystis Carinii* pneumonia, toxoplasmosis of the brain and Kaposi's sarcoma. The original definition of AIDS was revised in 1985 (CDC, 1985c) and again in 1987 (CDC, 1987j). Now there are plans to change the AIDS definition again to include all HIV-positive people with CD4⁺ cell count below 200; this would approximately double the number of people satisfying the AIDS definition.

1.3 HIV and AIDS in the United States

The history of AIDS in the United States (U.S.) is interesting. In June 1981 the Centers for Disease Control (CDC) reported on five homosexual men in Los Angeles with an uncommon *Pneumocystis carinii* pneumonia (PCP) who had a cellular immune dysfunction (CDC, 1981a). The next month CDC reported that 26 homosexual men in New York and California had Kaposi's sarcoma (KS), a rare skin cancer (CDC, 1981b). In August 1981 CDC reported more PCP and KS among homosexual men with immunosuppression (CDC, 1981c). Further study of some men with KS revealed that they had greatly reduced CD4⁺ T lymphocytes and that KS cases in these homosexual men were associated with a large number of sexual partners and a history of sexually transmitted diseases (Friedman-Klein et al., 1982). In 1982, the name *acquired immunodeficiency syndrome* (AIDS) was used to describe the people with this new condition. The clinical syndrome of AIDS was also recognized in heterosexual intravenous drug users (IVDUs) in the early 1980s (CDC, 1982). The retrovirus which causes AIDS was identified in 1983 and was called the *human T-lymphotropic virus* (HTLV-III) or *lymphadenopathy-associated virus* (LAV); it is now called the *human immunodeficiency virus* (HIV) or HIV-1 since an HIV-2 has been identified (CDC, 1986h). Thus important aspects of HIV transmission and AIDS in the U.S. were identified within a few years after the first cases were reported. Namely, that HIV was a virus which could be transmitted sexually and by needle-sharing IVDUs, that the immune system deteriorated due to the decline of CD4⁺ cells and that opportunistic infections and cancers occurred in the late stage of the disease.

Note that the *incidence* of a disease or condition is the number of new cases occurring per month (or year or other time interval), and that the *prevalence* is the total number of cases which exist at a given time. In an equilibrium situation for a disease such as measles, the prevalence is equal to the incidence times the average duration of the disease, but that this formula does not work for HIV or AIDS, since an equilibrium has not been reached and the duration of HIV infection is long and variable.

The yearly AIDS incidence in the U.S. has increased every year since 1980 and was about 44,000 in 1991 so that 22% of the 200,000 AIDS cases through 1991 occurred in 1991. The AIDS incidence is not uniform throughout the U.S. Figure 1.1 shows the 1991 annual case rates for the states. The incidence of AIDS is also not uniformly spread among people in the U.S., but is focused in certain risk groups and occurs more commonly in certain racial/ethnic groups.

So far, most of the AIDS cases in the U.S. have occurred in homosexual men with homosexual intercourse as the transmission mechanism. There have also been many AIDS cases in intravenous drug users (IVDUs) and some AIDS cases in homosexual men who are also IVDUs.

AIDS annual rates per 100,000 population, for cases reported in 1991, United States

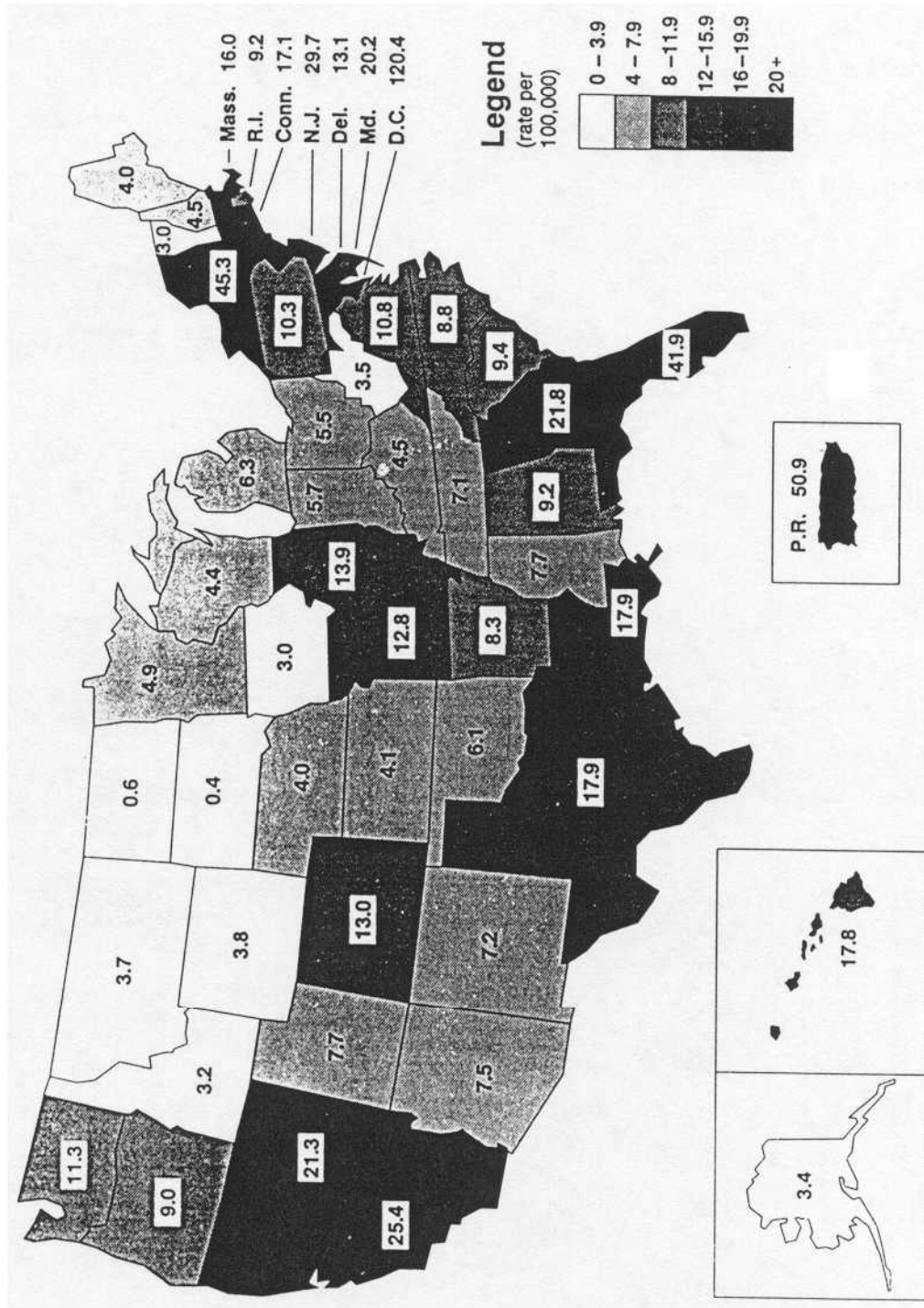


Figure 1.1. The geographical distribution of AIDS case rates in the United States. Source: CDC (1992, p. 3).

Most of the heterosexual AIDS cases in the U.S. are in the risk group of heterosexual partners of IVDUs. These four risk groups account for 94% of the AIDS cases in the U.S. in adults with a known risk of HIV infection. In children (<13 years old) 78% of AIDS cases with known risk of infection have occurred in children whose mother is an IVDU or a heterosexual partner of an IVDU (CDC, 1992). Since these are the major risk groups in the U.S., they are the groups analyzed by modeling in this monograph. Figure 1.2a shows the AIDS case rates by age group and sex in the U.S. The case rate in men is higher than in women because of homosexual and needle-sharing transmission (approximately three fourths of IVDUs with AIDS are men). Figure 1.2b shows the reported adult/adolescent AIDS cases by exposure category.

Blood-transfusion-acquired HIV infections are now very rare. Since 1985 all blood in the U.S. has been screened for HIV antibodies, with positive units being discarded. The risk group of being a recipient of a blood transfusion, blood component or tissue accounts for about 2% of all AIDS cases in the U.S., but this is primarily the result of those infected before screening started.

Many hemophiliacs in the U.S. were infected in the early 1980s by blood factor concentrate which was produced by pooling 5 to 10 thousand units of blood. Approximately three fourths of the 14,000 hemophiliacs in the U.S. have been infected with HIV. After HIV was identified, the blood factor concentrate was heat-treated to kill the HIV starting in late 1984. Hemophiliacs account for about 1% of AIDS cases in adults.

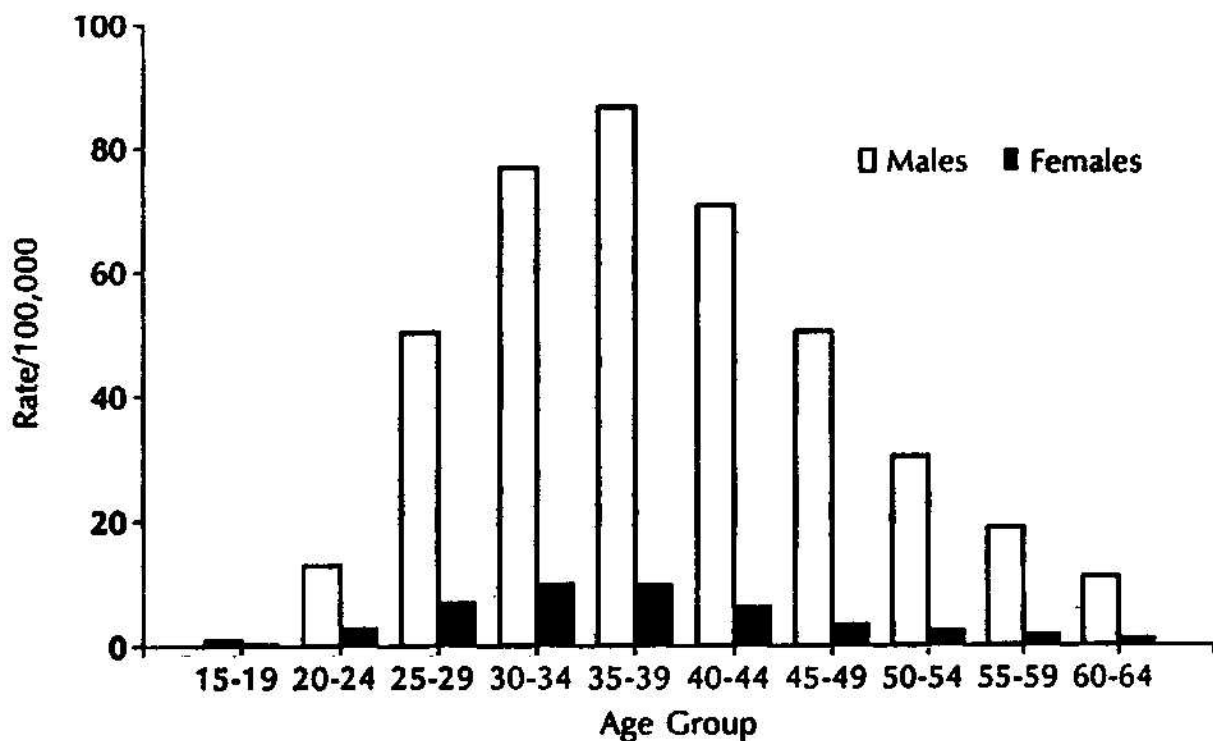
Other categories with small percentages of AIDS cases are heterosexual partners of people born in pattern II countries and children of women born in pattern II countries. Pattern II countries as described in the next section are those such as Haiti and many African countries, where heterosexual transmission predominates.

The primary source of data on AIDS in the U.S. is the Centers for Disease Control (CDC). All 50 states, the District of Columbia, U.S. dependencies and possessions and other related nations report AIDS cases to CDC using a uniform case definition and case report form. CDC compiles the data and reports it monthly in the publication, HIV/AIDS Surveillance Report. These reports include Tables of AIDS incidence in the states, large cities, adults, men, women, children, racial/ethnic groups and age groups. Tables 1.1, 1.2, and 1.3 are from the January 1992 HIV/AIDS Surveillance Report (CDC, 1992) which gives data through December 1991. These Tables give the numbers and percentages of AIDS cases in the various risk groups.

Note that 90% of AIDS cases in adults occur in men. Among white men, most AIDS cases occur in homosexual men, but among black and Hispanic men, the number of AIDS cases in homosexual men and in IVDUs are about equal. Among women, most cases occur in IVDUs, but many cases also occur in heterosexual partners of IVDUs. Observe that 29% of all AIDS cases occur in blacks, who constitute about 12% of the U.S. population, and 16% of all AIDS cases occur in Hispanics, who are about 8% of the U.S. population.

Reporting delays of AIDS cases are significant; about half of all AIDS cases are reported within 3 months of diagnosis, but about 15% are reported more than 1 year after diagnosis. Various methods are used for adjusting for reporting delays, but CDC uses a procedure developed by Karon, Devine and Morgan (1989). AIDS incidence data must be adjusted for reporting delays in order to see trends over time. Figure 1.3 shows AIDS incidences adjusted for reporting delays

ACQUIRED IMMUNODEFICIENCY SYNDROME (AIDS) – Annual rates per 100,000 adult population, by selected age group and sex for reported cases, United States, 1990



ACQUIRED IMMUNODEFICIENCY SYNDROME (AIDS) – Reported adult/adolescent cases, by exposure category, United States, 1990

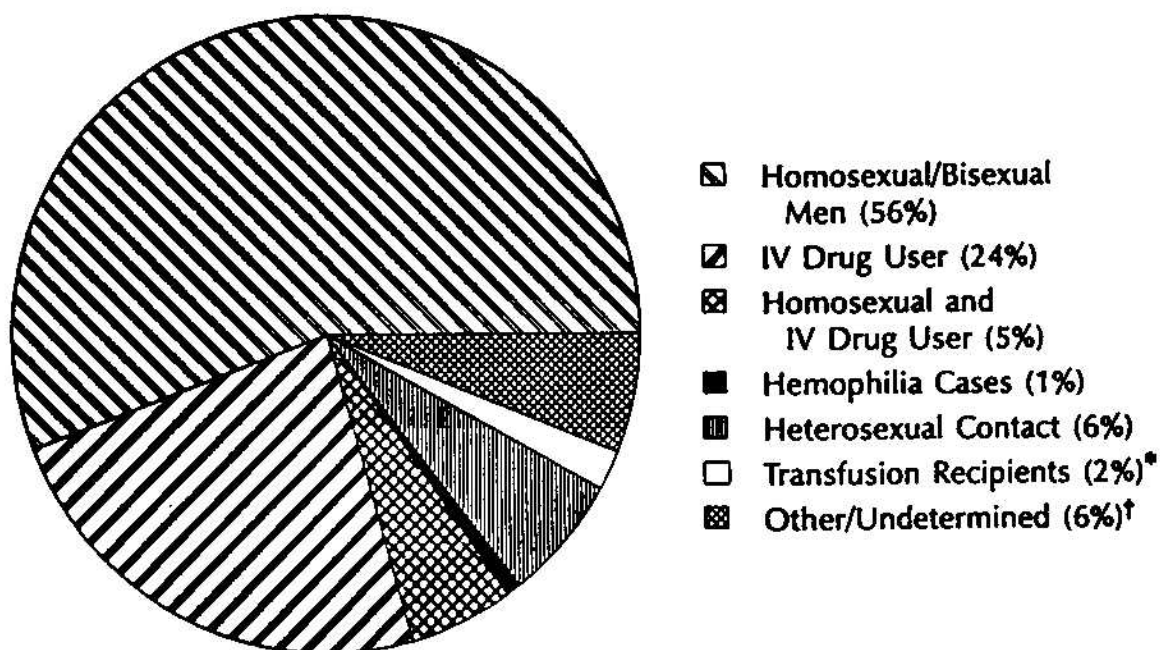


Figure 1.2 a) AIDS case rates by age and sex, b) AIDS case rates by exposure category.
Source: CDC (1990, p. 16).

Table 1.1 Male adult/adolescent AIDS cases by exposure category and race/ethnicity, reported through December 1991.
Source: CDC (1992, p. 11).

| Male exposure category | White, not Hispanic | | Black, not Hispanic | | Hispanic | | Asian/Pacific Islander | | American Indian/Alaska Native | | Total ⁴ | |
|--|---------------------|--------------|---------------------|-----------|-----------|---------------|------------------------|---------|-------------------------------|---------|--------------------|---------|
| | No. (%) | No. (%) | No. (%) | No. (%) | No. (%) | No. (%) | No. (%) | No. (%) | No. (%) | No. (%) | No. (%) | No. (%) |
| Men who have sex with men | 83,205 (80) | 20,540 (44) | 13,240 (46) | 936 (81) | 172 (64) | 118,362 (65) | | | | | | |
| Injecting drug use | 7,017 (7) | 16,798 (36) | 11,083 (39) | 40 (3) | 33 (12) | 35,048 (19) | | | | | | |
| Men who have sex with men and inject drugs | 7,547 (7) | 3,578 (8) | 1,925 (7) | 28 (2) | 41 (15) | 13,135 (7) | | | | | | |
| Hemophilia/coagulation disorder | 1,373 (1) | 127 (0) | 137 (0) | 18 (2) | 8 (3) | 1,671 (1) | | | | | | |
| Heterosexual contact: | 813 (1) | 3,307 (7) | 548 (2) | 9 (1) | 3 (1) | 4,687 (3) | | | | | | |
| Sex with injecting drug user | 483 | 1,077 | 313 | 5 | 3 | 1,882 | | | | | | |
| Sex with person with hemophilia | 7 | 1 | 2 | — | — | 10 | | | | | | |
| Born in Pattern-II ¹ country | 8 | 1,779 | 11 | 2 | — | 1,805 | | | | | | |
| Sex with person born in Pattern-II country | 40 | 48 | 9 | 1 | — | 98 | | | | | | |
| Sex with transfusion recipient with HIV infection | 43 | 19 | 16 | — | — | 79 | | | | | | |
| Sex with HIV-infected person, risk not specified | 232 | 383 | 197 | 1 | — | 813 | | | | | | |
| Receipt of blood transfusion, blood components, or tissue ² | 1,938 (2) | 418 (1) | 259 (1) | 55 (5) | 1 (0) | 2,679 (1) | | | | | | |
| Other/undetermined ³ | 2,287 (2) | 2,269 (5) | 1,432 (5) | 67 (6) | 11 (4) | 6,114 (3) | | | | | | |
| Male subtotal | 104,180 (100) | 47,037 (100) | 28,624 (100) | ,153(100) | 269 (100) | 181,696 (100) | | | | | | |

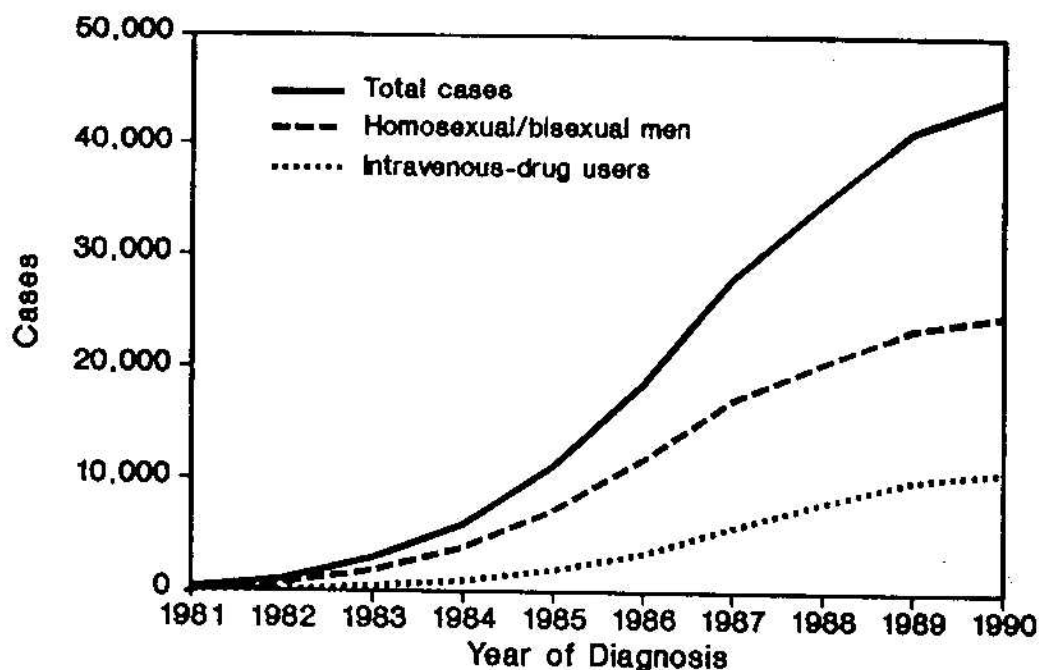
Table 1.2 Female adult/adolescent AIDS cases by exposure category and race/ethnicity, reported through December 1991.
Source: CDC (1992, p. 11).

| Female exposure category | 2,268 (41) | 6,185 (55) | 2,191 (50) | 16 (15) | 25 (56) | 10,705 (50) |
|---|-------------|--------------|-------------|-----------|----------|--------------|
| Injecting drug use | 29 (1) | 10 (0) | 3 (0) | — | — | 42 (0) |
| Hemophilia/coagulation disorder | 1,700 (31) | 3,784 (34) | 1,694 (39) | 37 (35) | 10 (22) | 7,249 (34) |
| Heterosexual contact: | | | | | | |
| Sex with injecting drug user | 859 | 2,244 | 1,343 | 14 | 7 | 4,484 |
| Sex with bisexual male | 343 | 216 | 78 | 11 | 1 | 651 |
| Sex with person with hemophilia | 78 | 10 | 4 | 2 | — | 94 |
| Born in Pattern-II country | 5 | 709 | 3 | — | — | 718 |
| Sex with person born in Pattern-II country | 10 | 63 | 2 | — | — | 76 |
| Sex with transfusion recipient with HIV infection | 102 | 28 | 28 | 2 | — | 161 |
| Sex with HIV-infected person, risk not specified | 303 | 514 | 236 | 8 | 2 | 1,065 |
| Receipt of blood transfusion, blood components, or tissue | 1,055 (19) | 349 (3) | 220 (5) | 35 (33) | 5 (11) | 1,668 (8) |
| Other/undetermined | 414 (8) | 828 (7) | 292 (7) | 17 (16) | 5 (11) | 1,561 (7) |
| Female subtotal | 5,466 (100) | 11,156 (100) | 4,400 (100) | 105 (100) | 45 (100) | 21,225 (100) |

Table 1.3 Pediatric (<13 years old) AIDS cases by exposure category and race/ethnicity, reported through December 1991.
Source: CDC (1992, p. 10).

| Pediatric (<13years old) exposure category | 112 (15) | 22 (1) | 26 (3) | 3 (18) | — | 163 (5) |
|---|-----------------|--------------------|-----------------|-----------------|---------------|-------------------|
| Hemophilia/coagulation disorder | | | | | | |
| Mother with/at risk for HIV infection: | 465 (63) | 1,704 (92) | 742 (87) | 8 (47) | 8(100) | 2,936 (85) |
| <i>Injecting drug use</i> | 224 | 833 | 365 | 2 | 2 | 1,430 |
| Sex with injecting drug user | 91 | 269 | 238 | 2 | 1 | 603 |
| Sex with bisexual male | 22 | 24 | 14 | 1 | — | 61 |
| Sex with person with hemophilia | 9 | 3 | 1 | — | — | 13 |
| Born in Pattern-II country | 1 | 241 | 2 | — | — | 244 |
| Sex with person born in Pattern-II country | — | 13 | — | — | — | 14 |
| Sex with transfusion recipient with HIV infection | 4 | 4 | 4 | — | — | 13 |
| Sex with HIV-infected person, risk not specified | 30 | 70 | 40 | 1 | 2 | 144 |
| Receipt of blood transfusion, blood components, or tissue | 22 | 25 | 13 | — | — | 60 |
| Has HIV infection, risk not specified | 62 | 222 | 65 | 2 | 3 | 354 |
| Receipt of blood transfusion, blood components, or tissue | 152 (21) | 65 (4) | 66 (8) | 6 (35) | — | 289 (8) |
| Undetermined | 10 (1) | 53 (3) | 20 (2) | — | — | 83 (2) |
| Pediatric subtotal | 739(100) | 1,844 (100) | 854(100) | 17 (100) | 8(100) | 3,471(100) |
| Total | 110,385 | 60,037 | 33,878 | 1,275 | 322 | 206,392 |

- a. Total cases, cases among homosexual/bisexual men[†], and cases among women and heterosexual men reporting intravenous (IV)-drug use



- b. Cases among persons reporting heterosexual contact with persons with, or at high risk for, HIV infection

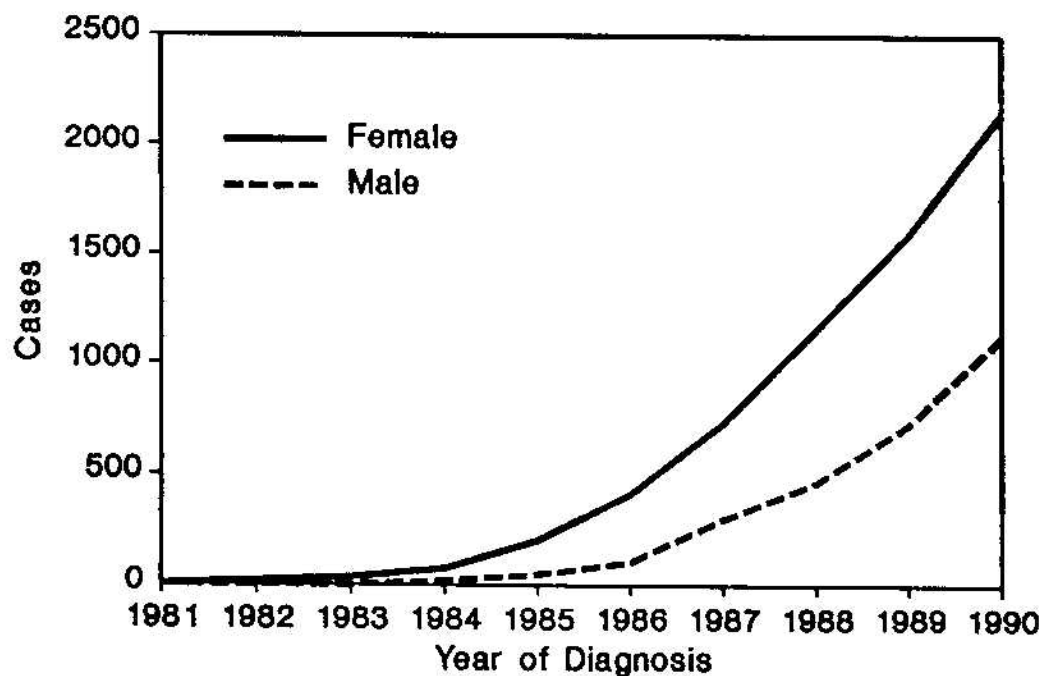


Figure 1.3. AIDS cases adjusted for reporting delays by year of diagnosis. a) total cases, homosexual/bisexual men, and IVDUs, b) heterosexual cases. Source: CDC (1991, p. 360).

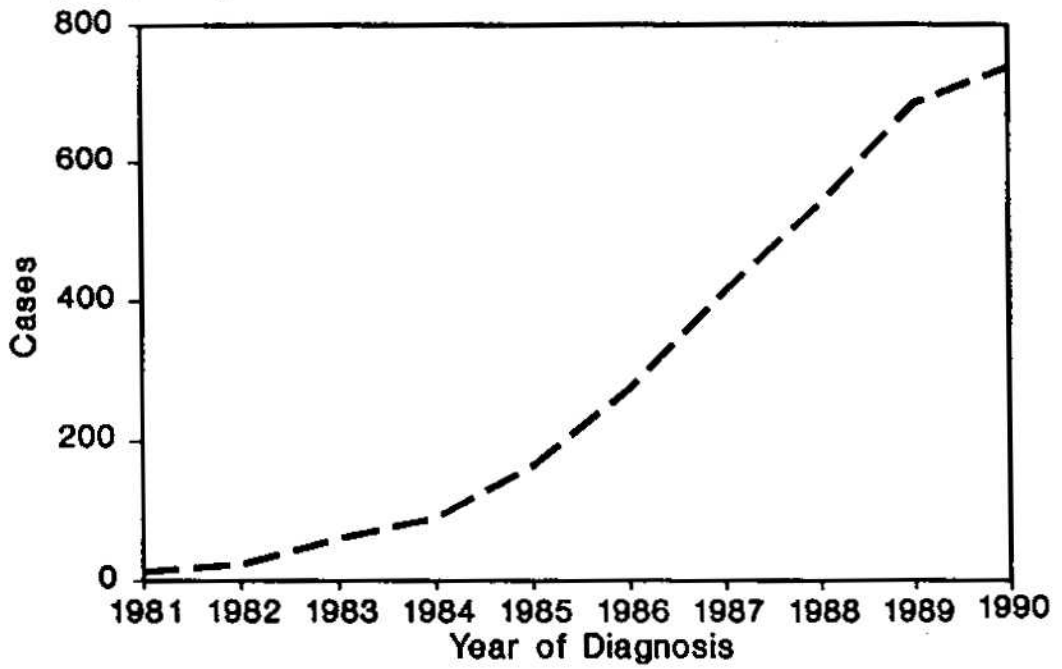
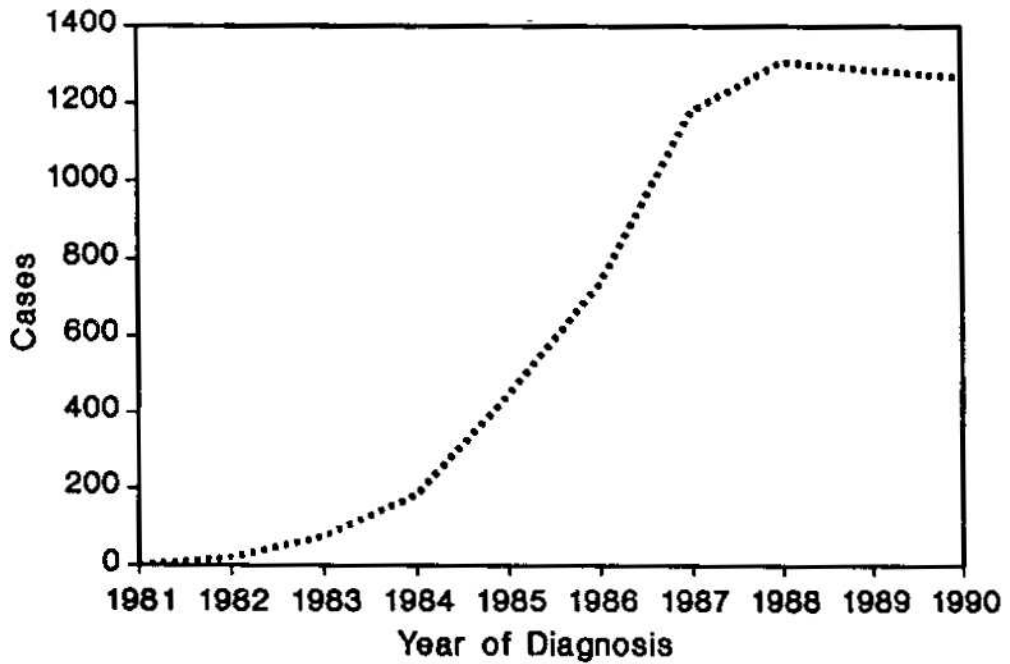
c. Perinatally acquired pediatric AIDS cases**d. Cases among recipients of transfusions of blood or blood products**

Figure 1.3. AIDS cases adjusted for reporting delays by year of diagnosis. c) perinatal cases, d) transfusion cases. Source: CDC (1991, p. 361).

for various risk groups (CDC, 1991b). From Figure 1.3a it appears that AIDS cases in homosexual men may be slowing down or leveling off. Also, AIDS incidence in IVDUs started later and continues to increase. In Figure 1.3b AIDS incidence in people reporting heterosexual contact started even later, but is now rising very rapidly. Pediatric AIDS cases in Figure 1.3c also continue to go up; the change in the last year may not be significant and due to reporting changes. The trend in recipients of transfusions of blood or blood products in Figure 1.3d is clear; AIDS cases in this group are now declining. This trend is not surprising since almost no new HIV infections occurred as a result of transfusions after blood screening started in early 1985.

Although national time trends in AIDS cases give some information, they often do not tell the whole story. Figure 1.4 shows that AIDS cases in homosexual men in New York City (NYC), San Francisco (SF), and Los Angeles (LA) appear to be leveling off after 1987, but AIDS cases in homosexual men in all other regions continue to increase at least through 1989 (Karon and Berkelman, 1991). Even this graph does not reveal the complete picture. Figure 1.5 shows AIDS incidence trends in NYC, LA, and SF; AIDS incidence appears to have started to decrease in NYC and has slowed down in LA and SF, but may not have leveled off yet in LA and SF. Karon and Berkelman (1991) also show that the trends in NYC, LA and SF are primarily due to changes in AIDS incidence in white homosexual men with smaller changes in blacks and Hispanics. In many other regions the AIDS incidence is increasing for both whites and non-whites. The take-home lesson from this analysis of AIDS in homosexual men is that it is often necessary to look beyond national data into regions, cities, and racial/ethnic groups to discover trends.

Data on HIV incidence and prevalence is relatively sparse and definitely less reliable than the AIDS incidence data. Some data on HIV incidence and prevalence is available from blood samples saved for other purposes; see Section 5.4.1 on San Francisco. But generally, there is no information on HIV incidence or prevalence in previous years. Even now it is very difficult to get information on HIV incidences or prevalences. Due to nonresponse bias and other factors, data from national surveys are often considered to be inaccurate. The widely-used value of 1 million HIV-infected people in the U.S. is based on a variety of estimates such as the back calculation method described in Section 1.5 (CDC, 1991b).

1.4 HIV and AIDS in the World

The AIDS pandemic is now entering its second decade. Although projections of its future are uncertain, it is clear that the world-wide epidemic is still in its early stages. In the 1990s millions of people who are already infected with HIV will develop AIDS and die. The World Health Organization (WHO) estimates that about 5 to 6 million men and 3 to 4 million women in the world have been infected with HIV. So far over 1 million have progressed to AIDS and an equal number have HIV-related illness. Worldwide, WHO forecasts that the cumulative number of HIV-infections will be 30 to 40 million by the year 2000. WHO projects that there will be 1 million adult AIDS cases and deaths per year, with about 1/2 million in Africa and 1/4 million in Asia (WHO, 1991).

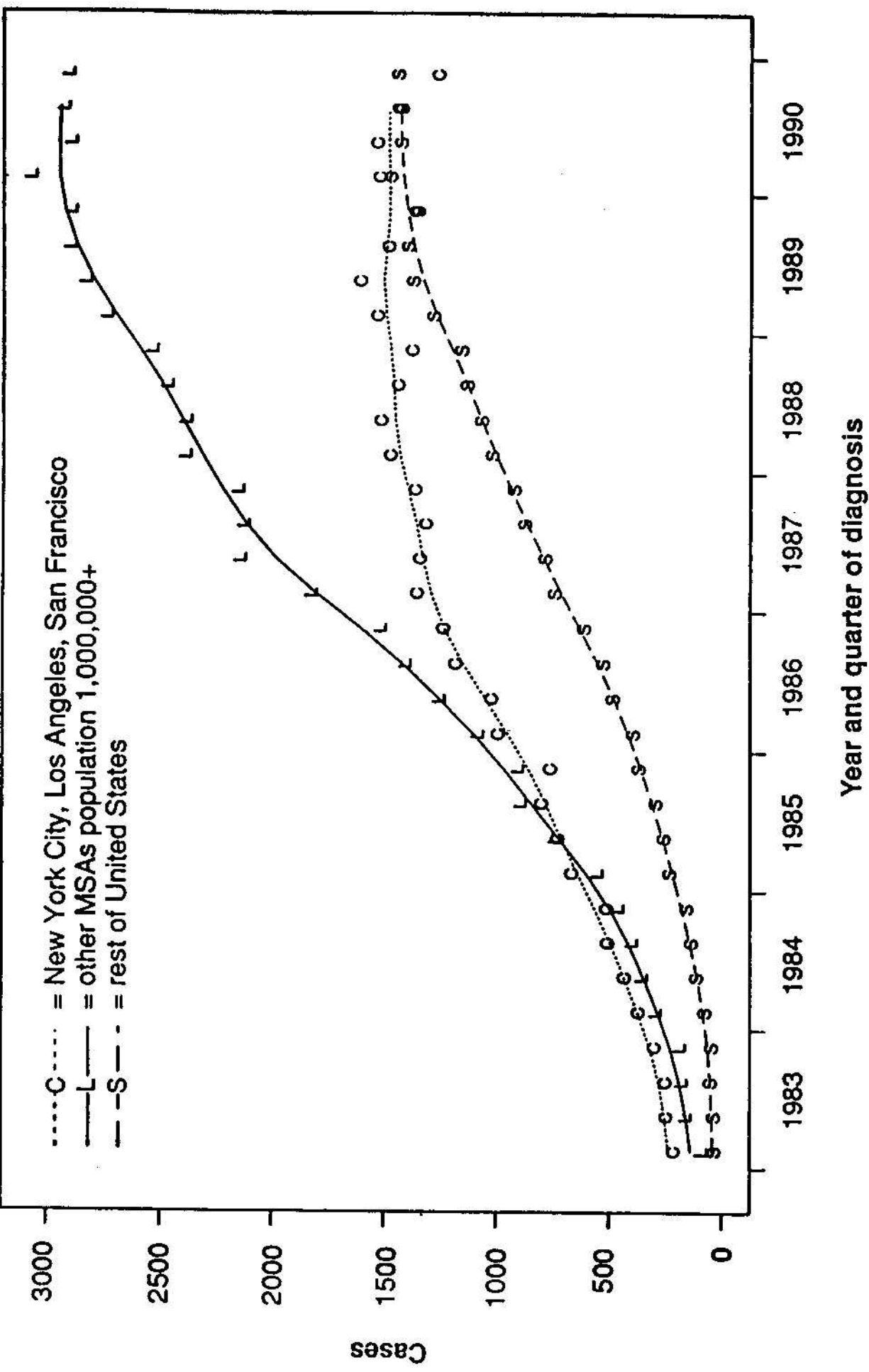
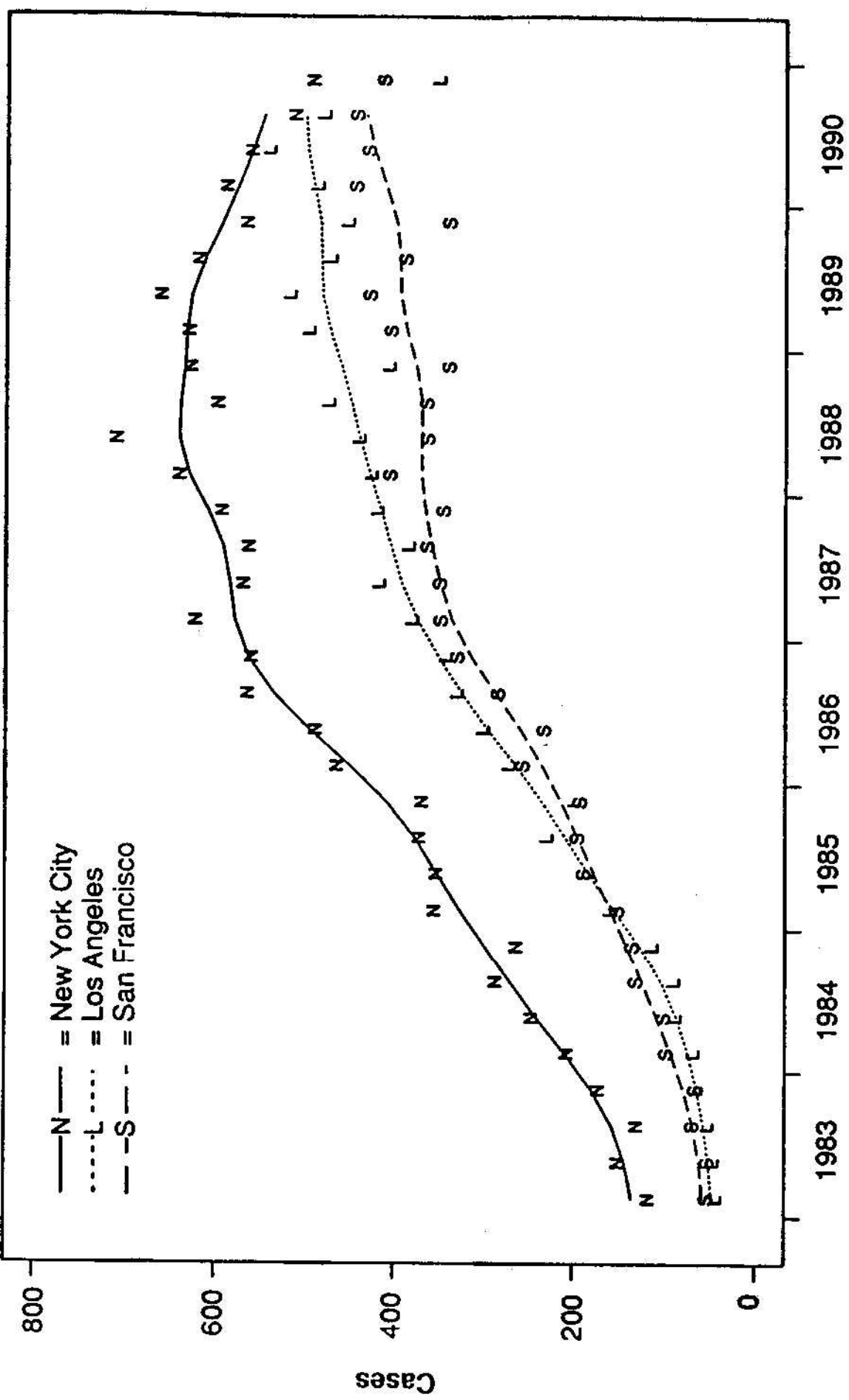


Figure 1.4. AIDS incidence adjusted for reporting delays in homosexual/bisexual men in the U.S. Source: Karon and Berkelman (1991).



Year and quarter of diagnosis

Figure 1.5. AIDS incidence adjusted for reporting delays in homosexual/bisexual men in three major cities (NYC, LA, SF) in the U.S. Source: Karon and Berkelman (1991).

The WHO has identified three basic patterns of HIV infection and AIDs in the world. Pattern I for AIDS occurs in North America, Western Europe, Australia and New Zealand. In these areas, the predominant modes of transmission are homosexual contact and needle-sharing by IVDUs, but there is some heterosexual and perinatal transmission. The male-to-female ratio is 10:1 to 15:1. There is no transmission by blood transfusion since blood is now screened. WHO suggests that the annual incidence of HIV infections in these areas may have peaked in the first half of the 1980s when there was rapid spread among homosexual men and IVDUs. The annual HIV incidences seem to have now decreased, presumably due to the reduction of risky behavior in these high risk groups. Heterosexual transmissions may increase, but WHO predicts that AIDS incidence will remain stable in the 1990s (WHO, 1991).

Pattern II for AIDS occurs in much of Africa and in parts of the Caribbean. In these areas the spread is primarily by heterosexual contact, and the male-to-female ratio is about 1:1. Some HIV transmission occurs during homosexual contact and by blood transfusion or medical uses of needles and syringes. Perinatal transmission is significant since there are many infected women. In sub-Saharan Africa the HIV epidemic grew in the 1970s and 1980s. The optimistic prediction of WHO is that the annual HIV incidence in this region will peak in the mid-1990s (WHO, 1991). Their prediction for Latin America, where the HIV incidence has been less, is similar, although there is the potential for high infection rates in Brazil.

Pattern III for AIDS occurs in the remainder of the world, including Asia, the Middle East and North Africa, where the HIV virus has arrived more recently. There is some HIV and AIDS in these areas; some HIV infections were acquired in other areas and some occurred within these areas. The AIDS incidence in these regions has been low so far, but there is the potential for increased incidence. In Asia, HIV transmission began in the late 1980s in a few countries, but HIV infection has spread rapidly since then. WHO predicts that annual HIV incidence in Asia will continue to rise until the early part of the next century, and by the end of the 1990s the annual incidence in Asia will exceed that in Africa (WHO, 1991).

Worldwide, heterosexual transmission is responsible for two-thirds of all HIV infections so far, and this fraction is expected to increase. About one out of three children born to an HIV-infected mother is HIV-infected and dies of AIDS, usually by age five. The other two children become orphans when their mother dies of AIDS. So far, there have been about 1 million HIV-infected infants born, over half of whom died of AIDS. There are now approximately 2 million uninfected children who are or will become orphans. WHO predicts that there will be 10 to 15 million AIDS orphans by the year 2000, most of them in sub-Saharan Africa (WHO, 1991).

The social and economic consequences of AIDS in developing countries will be stunning. Elderly people will not only be left without support as their grown children die, but also will become responsible for their surviving grandchildren. The costs of caring for people with HIV-related diseases will overwhelm the health care resources of many developing countries. There will be a decrease in the productivity of the workforce as young and middle-aged adults die from AIDS.

Meeting the challenge of AIDS requires extraordinary efforts throughout the world. Governments need to make AIDS prevention and control a top priority. They need to get all agencies and organizations involved in preventing HIV infections. They need to counter discrimination against HIV-infected persons and to fight complacency and denial about AIDS. The social conditions which put people at risk of HIV infection, such as multiple sexual partners, occur throughout the world. No racial and ethnic groups seem to have greater susceptibility or resistance to HIV infection. The prospects during the 1990s for developing a vaccine to prevent HIV infection or prevent AIDS are uncertain. Thus the primary means of prevention is through education of the people about the risks and possible preventative measures.

1.5 Three Approaches to Modeling HIV/AIDS

The objective of this section is to present descriptions and cite the literature on the main approaches to HIV/AIDS modeling. Surveys of mathematical models for HIV transmission and AIDS have been given by Isham (1988) and Anderson (1988b). An annotated bibliography of statistical methodology for study of HIV/AIDS has been published by Fusaro et al. (1989). Schwager, Castillo-Chavez and Hethcote (1989) have reviewed both statistical and mathematical approaches in HIV/AIDS modeling.

The first and most direct approach to predicting AIDS cases in the future is extrapolation (Morgan and Curran, 1986; Karon et al., 1988, 1989). This method is to fit an assumed form of the AIDS incidence curve to the AIDS incidence data in recent years and then to extend this curve for several years as a prediction of AIDS cases in the future. This method assumes that the current trends will continue for at least a few years into the future. Often separate curves and extrapolations are done for various risk groups. Advantages of extrapolation are its simplicity and ease of use. The extrapolation method has been a good predictor of AIDS incidence for a few years into the future, but it is not good for longer forecasts since it does not consider changes in the HIV epidemic due to factors such as behavioral changes or saturation in the high risk groups. Another disadvantage is that it does not give any information on HIV incidence or any understanding of the HIV transmission mechanisms. Extrapolation has been used on United Kingdom data by Healy and Tillett (1988) and on European data by Downs et al. (1987). Extrapolation has not worked as well in recent years because there have been clear changes in trends with a decreasing rate of growth and the incidence has reached a plateau in some risk groups.

The second approach to modeling AIDS incidence is usually called back calculation (Brookmeyer and Gail, 1986, 1988; Gail and Brookmeyer, 1988; Brookmeyer and Damiano, 1989). The total number of cases of AIDS at time t is the summation up to time t of the product of the HIV incidence at time τ and the probability of developing AIDS within $t - \tau$ years after infection. Thus if the HIV incidence and the distribution of the AIDS incubation period were known up to time t , then the cumulative AIDS cases would be calculated in a straight-forward way using the convolution summation above. Back calculation is a deconvolution process; it uses a given AIDS incidence up to time t and an estimated distribution for the AIDS incubation period to estimate the HIV incidence up to time t . This HIV incidence up to time t and its

extrapolation for a few years are then used to forecast the AIDS incidence for a few years. The distribution of the incubation period for AIDS can be estimated parametrically or nonparametrically. The back calculation procedure is often applied to separate risk groups.

Back calculation has been used for forecasting AIDS incidence for a few years and does have the advantage that it also yields estimates of HIV incidence; however, there are several disadvantages. It does not yield any information on the HIV transmission dynamics or estimates of parameter values. Estimated distributions for the AIDS incubation period are uncertain and the back calculation procedure is very sensitive to the distribution used (Brookmeyer and Damiano, 1988; Hyman and Stanley, 1988). The instability of the back calculation process implies that the confidence intervals of the estimates of HIV incidence and future AIDS incidence are very wide.

The third approach to modeling AIDS is to use HIV transmission dynamics models which include the progression to AIDS. These models often have the population divided into compartments consisting of those who are susceptible, in each of the infectious stages, or in the AIDS phase. In deterministic transmission models, the movements between these compartments by becoming infected, progressing to the next stage or AIDS, migrating or dying are specified by systems of difference or differential equations. The models of this type used in this monograph are formulated in Chapters 3 and 7. Dynamic models and computer simulations are experimental tools for comparing regions or risk groups, testing theories, assessing quantitative conjectures and answering questions. For example, in Chapter 6 we determine the relative importance of saturation in the high risk group and behavior changes in changing the AIDS incidence in homosexual men in San Francisco. Modeling can also be used to theoretically evaluate, compare or optimize various detection, prevention, intervention or control programs. A detailed discussion of the purposes and limitations of dynamic epidemiological modeling is given in the next section.

Some HIV transmission dynamics models are stochastic with probabilities of moving to the next stage at each time step. Mode et al. (1989) investigated the relationship between Monte Carlo simulations of a stochastic HIV/AIDS model and solutions of a deterministic model using the expected values. Tan (1991) looks at general stochastic models for the simulation of HIV transmission and AIDS. In the book (Tan, 1992) the stochastic analog of the deterministic model used in Chapters 5 and 6 is analyzed and the results are compared with those for the deterministic model. A special issue (vol 107, No 2, December 1991) of the journal *Mathematical Biosciences* contains six papers on stochastic models for HIV/AIDS.

Both deterministic and stochastic dynamic epidemiological models have been used to analyze many diseases including measles, rubella, mumps, chicken pox, poliomyelitis, smallpox, gonorrhea, schistosomiasis, malaria, cholera and rabies. For examples, see the books by Bailey (1975), Anderson (1982), Anderson and May (1982, 1991) and Hethcote and Yorke (1984). Many dynamic models have been analyzed mathematically; see Hethcote (1989a) for an introduction to mathematical epidemiology or Hethcote et al. (1981) and Hethcote and Levin (1989) for surveys of mathematical results.

In recent years there has been a tremendous number of modeling papers which use dynamic models of HIV transmission and progression to AIDS. Surveys of these modeling papers are cited

at the beginning of this section. Two authors who have contributed to HIV/AIDS modeling are Anderson and May (e.g., Anderson, 1988; Anderson and May, 1988; May, 1988); some of their modeling results are incorporated into their encyclopedic book (Anderson and May, 1991), which is built from their numerous epidemiological papers during the past ten years. Their models have covered not only homosexual and heterosexual populations in the United States and United Kingdom, but also heterosexual populations in Africa. They have estimated reproduction numbers (contact numbers), doubling times and demographic consequences.

Some of the multigroup models for HIV/AIDS are similar in formulation to the multigroup models developed for gonorrhea, which is another sexually transmitted disease (STD) (Lajmanovich and Yorke, 1976; Hethcote and Yorke, 1984). Multigroup models are appropriate for STDs because people are heterogeneous in their behaviors and contact rates with others. The contacts can be homosexual, heterosexual or needle-sharing contacts. The models in this monograph in Chapters 3 and 7 are multigroup models.

Proportionate mixing in multigroup models was used in gonorrhea modeling by Nold (1980) and Hethcote and Yorke (1984). They also used a convex combination of internal contacts and proportionate mixing, which has been named *preferred mixing* by Jacquez et al. (1988) and *biased mixing* by Hyman and Stanley (1988, 1989). Jacquez et al. (1988) used multigroup compartmental models for HIV with constant recruitment into susceptible classes and variable infectivity in the infectious stages to analyze the effects of different mixing patterns. This analysis has been extended to more complicated mixing structures by Jacquez et al. (1989) and Koopman et al. (1989). Castillo-Chavez and coauthors (Blythe and Castillo-Chavez, 1989; Busenberg and Castillo-Chavez, 1989; Castillo-Chavez et al., 1989) have focused on the theoretical analysis of diverse mixing structures in HIV/AIDS models.

Hyman and Stanley (1988, 1989) have considered both continuous and discrete HIV/AIDS models with heterogeneity and different mixing structures. They have analyzed the spread from high to low risk groups, the effects of variable infectivity and the instability of the back calculation procedure. An innovative approach to dynamic modeling of HIV/AIDS has been developed by Dietz and Haderler (Dietz, 1988; Dietz and Haderler, 1988; Haderler et al., 1988). Their models explicitly consider the dynamics of pairs of individuals and the duration of their partnerships. These pair formation and dissolution models are distinctly different from the *contact rate mixing matrix* models used by many other modelers.

We have not attempted to give a complete survey of HIV/AIDS modeling; other papers are cited in later chapters. For more information the reader is referred to the survey articles cited earlier or the collection of papers edited by Castillo-Chavez (1989). The primary features which distinguish the approach in this monograph are the use of the deterministic dynamic simulation models as a framework for the interpretation of data, the efforts to estimate parameter values from data, the fitting of the model to HIV and AIDS incidence data and comparisons obtained by fitting the same model in fifteen different subregions of the U.S.

1.6 Purposes and Limitations of Epidemiological Modeling

Epidemiology is the study of the distribution and determinants of disease prevalence in humans. Epidemiologists study both infectious diseases and chronic diseases such as cancer and cardiovascular disease. One function of epidemiology is to describe the distribution of the disease, i.e., find out who has how much of what, where and when. Another function is to identify the causes or risk factors for diseases in order to find out why everyone doesn't have the same thing uniformly. A third function of epidemiology is to build and test theories. A fourth function is to plan, implement and evaluate detection, control and prevention programs.

Epidemiological modeling can play an important role in the last two functions above. Here we focus on modeling infectious diseases in human populations and do not consider models for chronic diseases such as cancer. In most of this section "epidemiological modeling" refers to dynamic modeling where the population is divided into compartments based on their health status (e.g. susceptible, infectious, recovered) and the movements between compartments by becoming infected, progressing, recovering or migrating are specified by differential or difference equations. First, fifteen purposes of epidemiological modeling are described and then three limitations are considered; they are all summarized in Table 1.4. One cannot infer from the numbers of purposes and limitations that the advantages overwhelm the limitations since the first limitation is very broad.

A primary reason for infectious disease modeling is that it leads to clear statements of the assumptions about the biological and social mechanisms which influence disease spread. The model formulation process is valuable to epidemiologists and modelers because it forces them to be precise about the relevant aspects of transmission, infectivity, recovery and renewal of susceptibles. Epidemiological modelers need to formulate models clearly and precisely using parameters which have well-understood epidemiological interpretations such as a contact rate or an average duration of infection (Hethcote, 1976, 1989a). Complete statements of assumptions are crucial so that the reasonableness of the model can be evaluated in the process of assessing the conclusions.

An advantage of mathematical modeling of infectious diseases is the economy, clarity and precision of a mathematical formulation. A model using difference, differential, integral or functional differential equations is not ambiguous or vague. Of course, the parameters must be defined precisely and each term in the equations must be explained in terms of mechanisms, but the resulting model is a definitive statement of the basic principles involved. Once the mathematical formulation is complete, there are many mathematical techniques available for determining the threshold, equilibrium, periodic solutions, and their local and global stability. Thus the full power of mathematics is available for analysis of the equations. Moreover, information about the model can also be obtained by numerical simulation on digital computers of the equations in the model. The mathematical analyses and computer simulations can identify important combinations of parameters and essential aspects or variables in the model. In order to choose and use epidemiological modeling effectively on specific diseases, one must understand the behavior of the available formulations and the implications of choosing a particular formulation.

Table 1.4. Purposes and Limitations of Epidemiological Modeling**Fifteen Purposes of Epidemiological Modeling**

1. The model formulation process clarifies assumptions, variables and parameters.
2. The behavior of precise mathematical models can be analyzed using mathematical methods and computer simulations.
3. Modeling allows explorations of the effect of different assumptions and formulations.
4. Modeling provides concepts such as a threshold, reproduction number, etc.
5. Modeling is an experimental tool for testing theories and assessing quantitative conjectures.
6. Models with appropriate complexity can be constructed to answer specific questions.
7. Modeling can be used to estimate key parameters by fitting data.
8. Models provide structures for organizing, coalescing and cross-checking diverse pieces of information.
9. Models can be used in comparing diseases of different types or at different times or in different populations.
10. Models can be used to theoretically evaluate, compare or optimize various detection, prevention, therapy and control programs.
11. Models can be used to assess the sensitivity of results to changes in parameter values.
12. Modeling can suggest crucial data which needs to be collected.
13. Modeling can contribute to the design and analysis of epidemiological surveys.
14. Models can be used to identify trends, make general forecasts, or estimate the uncertainty in forecasts.
15. The validity and robustness of modeling results can be assessed by using ranges of parameter values in many different models.

Three Limitations of Epidemiological Modeling

1. An epidemiological model is not reality; it is an extreme simplification of reality.
2. Deterministic models do not reflect the role of chance in disease spread and do not provide confidence intervals on results.
3. Stochastic models incorporate chance, but are usually harder to analyze than the corresponding deterministic model.

Thus mathematical epidemiology provides a foundation for the applications (Hethcote et al., 1981; Hethcote and Levin, 1989).

Modelers are able to explore and examine the effects of different assumptions and formulations. For example, they can compare an ordinary differential equations model, which corresponds to a negative exponential distribution for the infectious period, with a delay-differential equations model, which assumes that everyone has the same fixed, constant-length infectious period. They can examine the impact of assuming homogeneous mixing instead of heterogeneous mixing. They can study the influence of different mixing patterns between groups on the spread of a disease in a heterogeneous population (see Jacquez et al., 1988; Castillo-Chavez et al., 1989). They can decide if models with and without an exposed class for people in the latent period behave differently. The results of exploring different formulations may provide insights for epidemiologists and are certainly useful to modelers who are choosing models for specific diseases. For example, if the essential behaviors (thresholds and asymptotic behaviors) are the same for a formulation using ordinary differential equations and a formulation using delay differential equations, then the modeler would probably choose to use the ordinary differential equations model since it is simpler (Hethcote and Tudor, 1980).

Epidemiologists need to understand the concept of a threshold which determines whether the disease persists or dies out. They need to know that the reproduction number is a combination of parameters which gives the number of secondary cases "reproduced" by a typical infective during the infectious period in a population where everyone is susceptible. They need to be aware that this reproductive number is the same as the contact number, which is the average number of adequate contacts of an infective person during the infectious period (Hethcote and Van Ark, 1987). They need to realize that the average replacement number is one at an endemic equilibrium because, in an average sense, each infective replaces itself by one new infective in an equilibrium situation. Probably the most valuable contributions of mathematical modeling to epidemiologists are concepts like those above. Mathematical models serve as a framework to explain causes, relationships and ideas. Mathematical models are very useful in obtaining conclusions that have an easily understood interpretation. Helping epidemiologists internalize conceptual and intuitive understandings of disease processes, mechanisms and modeling results is an essential aspect of the work of mathematical modelers.

Epidemiological modeling is an important part of the epidemiologist's function to build and test theories. Mathematical and computer simulation models are the fundamental *experimental tools* in epidemiology. Experiments with infectious diseases in actual populations are often unethical or very expensive or impractical. Thus modeling is essential for exploratory work. The only data usually available are from naturally occurring epidemics or from the natural incidence of endemic diseases; unfortunately, even these data are not complete since many cases are not reported. Since repeatable experiments and accurate data are usually not available in epidemiology, mathematical and computer simulation models must be used to perform needed theoretical experiments with different parameter values and different data sets. It is easy in a computer simulation to find out what happens if one or several parameters are changed.

Another advantage of epidemiological modeling is the availability of models to assess quantitative conjectures. For example, models can be used to check the claim that a two-dose vaccination program for measles would lead to herd immunity whereas a one-dose program would not (Hethcote, 1983). With a model one could see if the AIDS virus would eventually disappear in a population of homosexual men if half of them became celibate or practiced safer sex. Although it is understood how an infective can transmit the infection by contacts with others and how a disease spreads through a chain of infections, it is not easy to comprehend the large scale dynamics of disease spread in a community or to evaluate quantitative theories without the formal structure of a mathematical model. An epidemiological model connects the microscopic description or role of an infectious individual to the macroscopic spread of the infection in a community. Mathematical models are a formal, quantitative way of building and testing theories.

When formulating a model for a particular disease, it is necessary to decide which factors to include and which to omit. This choice will often depend on the particular question that is to be answered. Simple models have the advantage that there are only a few parameters, but have the disadvantage of possibly being naive and unrealistic. Complex models may be more realistic, but they may contain many parameter values for which estimates cannot be obtained. The art of epidemiological modeling is to make suitable choices in the model formulation so that it is as simple as possible and yet is adequate for the question being considered. It is important to recognize both the capabilities and limitations of epidemiological modeling. Many important questions cannot be answered using a given class of models. The most difficult problem for a modeler is to find the right combination of available data, an interesting question and a mathematical model which can lead to the answer.

For AIDS one can ask whether needle-sharing is more important for transmission than sexual activity. How important are homosexual intravenous drug users in linking the epidemics in the homosexual/bisexual men and intravenous drug users? Is there a "core" group for AIDS as there was for gonorrhea and is it the same core group? How important are bisexuals and prostitutes in the transmission process? How rapidly does HIV infection spread from high risk groups to low risk groups? Why do AIDS cases in many populations increase initially as a cubic function of time? Hyman and Stanley (1988) suggest that this cubic growth could be explained by the spread of infection from high contact-rate groups down through lower contact-rate groups. Is sustained transmission likely in the heterosexual population with frequent partner changes? What effect will therapy with zidovudine (AZT) have on the average survival time with AIDS and on the transmission? There are many relevant questions to ask about AIDS.

By fitting the incidence predicted by a model to actual incidence data (number of new cases per month), it is possible to estimate key parameters such as contact numbers or average durations. Contact numbers (reproduction numbers) are estimated for various directly-transmitted diseases in Anderson and May (1982, 1991). Hethcote and Van Ark (1987) present methods for estimating contact numbers for the groups in a multigroup model for disease transmission.

In AIDS modeling, fitting the predicted HIV prevalence and AIDS incidence to data on both can lead to estimates of the average number of new partners per month and the probability

of HIV transmission per new partner (see Chapter 6). Since these parameters have already been estimated from the literature, the modeling provides a check on these *a priori* estimates. Thus the model is a way of coalescing parameter estimates and other pieces of data to see if they fit into a consistent framework. The model can serve as the conceptual and quantitative unifier of all the available information.

Comparisons can lead to a better understanding of the process of disease spread. It may be enlightening to compare the same disease at different times, the same disease in different populations or different diseases in the same population. One method for comparing diseases is to estimate parameter values for the diseases and then compare the parameter values. For example, in Anderson and May (1982, 1991) and Hethcote (1989a) estimates of contact numbers (also called reproduction numbers) are given for various diseases such as measles, whooping cough, chicken pox, diphtheria, mumps, rubella, poliomyelitis and smallpox. These numbers lead to fractions of the population which must be immune in order to achieve herd immunity for these diseases. Childhood diseases are compared in rural and urban areas in Anderson (1982). The level of HIV infection in homosexual men in San Francisco is compared at various times and the changes in behavior in this population are given in Chapter 6.

A very important reason for epidemiological modeling is the value of models for theoretical evaluations and comparisons of detection, prevention, therapy and control programs. Epidemiologists and politicians need to understand the effects of different policy decisions. Qualitative predictions of infectious disease models are always subject to some uncertainty since the models are idealized and the parameter values can only be estimated. However, quantifications of the relative merits of several control methods or of control versus no control are often robust in the sense that the same conclusions hold for a broad range of parameter values and a variety of models. Control strategies for gonorrhea such as screening, rescreening, infectee tracing, infector tracing and potential vaccines for both women and men were compared in Hethcote and Yorke (1984). Various vaccination strategies for rubella and measles have been compared using modeling (Hethcote, 1983, 1989b; Anderson and May, 1982).

Sometimes an optimization method can be used on some aspect of an epidemiological model. Hethcote and Waltman (1973) used dynamic programming to find an optimal vaccination strategy to control an epidemic at least cost. Longini et al. (1978) determined the best age or social groups to vaccinate for Hong Kong and Asian influenza when there was a limited amount of vaccine available. Hethcote (1988) found theoretical optimal ages of vaccinations for measles in three geographic locations. See Wickwire (1977) for a survey of modeling papers on the control of infectious diseases.

Predictions or quantitative conclusions of a model are said to be sensitive to a parameter if a small change in the parameter causes a large change in the outcome. A model is insensitive to a parameter if the outcomes are the same for a wide range of values of that parameter. An important reason for epidemiological modeling is the determination of the sensitive and insensitive parameters since this information provides insight into the epidemiological processes. Once the sensitive parameters are identified, it may be possible to make special efforts to gather data to obtain better estimates of these parameters. Thus the modeling process can help identify crucial

data that should be collected. If the data and information are insufficient to estimate parameters in a model or to piece everything together, then the model formulation process can identify the missing information which must be gathered.

Modeling can enhance an epidemiological survey by providing a framework for the analysis of the survey results. In the design phase modeling can help identify important questions that need to be asked in order to have adequate data to obtain results. The statistical aspects of a survey design are now analyzed before the survey is conducted. Similarly, those design aspects relevant to parameter estimation in models should be analyzed before the survey is conducted to see if all significant data are being collected.

Another purpose of epidemiological modeling is to make forecasts about the future incidence of a disease. Although people often think of prediction as the first or only purpose of epidemiological modeling, the purposes given above are more important. Accurate forecasts are usually not possible because of the idealization in the models and the uncertainty in the parameter values. However, possible forecasts under various scenarios can sometimes be given or trends can be identified. One purpose of epidemiological modeling is to reduce the uncertainty in future incidence projections. For example, improved estimates of the number of HIV infections and AIDS cases in future years are needed in order to provide appropriate public health programs and adequate medical care for patients with HIV-related medical conditions and AIDS. Although simple extrapolations for 2 or 3 years of the time trends for numbers with AIDS or in a final stage of HIV infection are often adequate, better estimates are obtained with models which consider population changes due to the disease transmission process.

The validation of models and modeling results is difficult since there are rarely enough good data to adequately test or compare different models with data (Anderson and May, 1991). However, when the same concept or result is obtained for a variety of models and by several modelers, then the confidence in the validity of the result is increased. For example, three different papers using different approaches to comparing two rubella vaccination strategies agree on the criteria determining situations when each strategy is best (Hethcote, 1989b). Similarly, when a quantitative result holds for a wide range of parameter values, then this result has some robustness. For example, if the same relative values of a variety of control procedures are obtained for wide ranges of parameter values, then the relative values may be robust even though the absolute values may vary (see Hethcote and Yorke, 1984).

After discussing the purposes and advantages of epidemiological modeling, it is imperative to also discuss the limitations. Epidemiologists and policy makers need to be aware of both the strengths and weaknesses of the epidemiological modeling approach.

The first and most obvious limitation is that all epidemiological models are simplifications of reality. For example, it is often assumed that the population is uniform and homogeneously mixing; this is always a simplification, but the deviation of reality from this simplification varies with the disease and the circumstances. This deviation from reality is rarely testable or measurable; however, it can sometimes be estimated intuitively from an understanding of the disease epidemiology. As described in the introduction, part of the reason for the differences between the model's behavior and the actual spread of the disease is that transmission is based on

the interactions of human beings, and homo sapiens have highly variable behavior and interactions. People do not behave in reasonably predictable ways like molecules or cells or particles.

Because transmission models are simplifications with usually unknown relationships to actual disease spread, one can never be completely certain about the validity of findings obtained from modeling such as conceptual results, experimental results, answers to questions, comparisons, sensitivity results and forecasts. Even when models are made more complicated in order to better approximate actual disease transmission, they are still abstractions. For example, multigroup models based on risk groups or sexual activity levels or age structure still assume that the distinctions between groups are clear and sharp and that the mixing between groups follows some assumed pattern. The modeler must always exercise judgment in deciding which factors are relevant and which are not when analyzing a specific disease or question.

The modeler needs to know that an SIR model is usually adequate since the behavior is essentially the same for the SEIR model where E is the latent period class. The notation SEIR for a model means that susceptibles (S) move to an exposed class (E) (when they are in a latent period; i.e., they are infected, but not yet infectious), then move to an infectious class (I), and finally to a removed class (R) when they recover and have permanent immunity. The modeler must realize that assuming that the infectivity is constant during the infectious period is reasonable for diseases like measles with a short infectious period (about 8 days), but is not reasonable for a disease like the human immunodeficiency virus (HIV) where the average time from HIV infection to AIDS is about 10 years. To incorporate age structure, the modeler must decide whether to use a model with time and age as continuous variables or to use discrete compartments corresponding to age brackets; the former requires the estimation of birth, death and aging rate coefficients as functions of time while the latter requires the estimation of birth, death and transfer rate constants. The modeler needs to decide whether the disease is epidemic in a short time period so that vital dynamics (births and deaths) can be excluded or the disease is endemic over a long time period (many years) so that vital dynamics must be included. The modeler must be aware that as the complexity of a model is increased so that it better approximates reality, the number of parameters increases so that it becomes increasingly difficult to estimate values of all of these parameters. Thus the modeler must (consciously or unconsciously) make many decisions regarding the relevant aspects in choosing a model for a specific disease or question.

When describing the purposes for epidemiological modeling, we have been thinking primarily of deterministic models, but the purposes also apply to stochastic epidemiological models. Deterministic models are those which use difference, differential, integral or functional differential equations to describe the changes in time of the sizes of the epidemiological classes. Given the starting conditions for a well-posed deterministic model, the solutions as a function of time are unique. In stochastic models, there are probabilities at each time step of moving from one epidemiological class to another. When these models are simulated with the probabilities calculated using random number generators, the outcomes of different runs are different so that this approach is often called Monte Carlo simulation; conclusions are obtained by averaging the

results of many computer simulations. Simple deterministic models for epidemics have a precise threshold which determines whether an epidemic will or will not occur. In contrast, stochastic models for epidemics yield quantities such as the probability that an epidemic will occur and the mean time to extinction of the disease. Thus the approach, concepts and appropriate questions are quite different for stochastic models (Bartlett, 1960; Ludwig, 1974; Bailey, 1975; Nasell, 1985; Becker, 1989).

Both deterministic and stochastic epidemiological models have other limitations besides being only approximations of reality. Deterministic models do not reflect the role of chance in disease spread. Sometimes parameter values in deterministic models are set equal to the mean of observed values and the information on the variance is ignored. A set of initial conditions leads to exactly one solution in a deterministic model; thus no information is available on the reliability of or the confidence in the results. When quantities such as the contact number, replacement number, average infectious period, etc. are estimated by fitting output to data, confidence intervals on these estimates are not obtained. Confidence intervals also can not be found for comparisons and forecasts. Some understanding of the dependence on parameter values is obtained through a sensitivity analysis where the effects of changes in parameter values on results are found. If the variance of the observed value of a parameter is high and the results are sensitive to that parameter, then the confidence in the results would be low. However, the lack of precise confidence intervals implies that the reliability of the results is uncertain.

Stochastic epidemiological models incorporate chance, but it is usually harder to get analytic results for these models. Moreover, computational results are also harder since Monte Carlo simulations require many computer runs (25 to 100 or more) in order to detect patterns and get quantitative results. Even for stochastic epidemiological models where parameters are estimated by fitting the mean of the simulations to data, it may not be possible to find confidence intervals on these parameter estimates.

1.7 Expectations of Modeling HIV Transmission Dynamics and AIDS

In June 1988 the Presidential Commission on the HIV Epidemic recommended that "the research community should be actively engaged in developing innovative models that better describe and explain the transmission of HIV within the population." On July 25-29, 1988 the AIDS Modeling and Epidemiology Workshop was held in Leesburg, Virginia. Ninety scientists in the fields of mathematics, statistics, epidemiology, data management, biology, behavioral sciences and social sciences participated. "The purpose of the Workshop was to examine the current status of AIDS and HIV modeling, to assess the potential benefits from mathematical and statistical analysis, and to make recommendations for a program of research." The conclusions from this meeting given below are from the report of the Office of Science and Technology Policy, Executive Office of the President (OSTP, 1988).

"The major findings of the workshop are

1. "Mathematical modeling and statistical analysis must be brought to bear fully on the AIDS epidemic. The success of

mathematical techniques in other fields of application and the presence of modern computational capability argue strongly for this added approach to AIDS and HIV analysis and projection. The modeling of the gonorrhea epidemic earlier this decade, cost-benefit analysis of rubella in England, and predictions of the spread of Hong Kong flu have exhibited that modeling and statistical projections of epidemics can make key contributions."

2. "Mathematical and statistical methods that have been used for sexually transmitted diseases and other epidemics are only partially applicable to AIDS and HIV. The basic structure will be similar but empirical parameters and functions that characterize the new disease are required: intra- and inter-risk group mixing, partner relationships and duration, and biological transmissibility and infectivity during the period between HIV and AIDS. Progress has been made in formulating these parameters, but more data are required and a diverse set of modeling techniques and models are appropriate."

3. "Given a supply of biological and behavioral data of good quality, modeling and advanced statistical analysis can provide better estimates of AIDS and HIV prevalence than are now available and eventually of HIV incidence. Cost-benefit assessments of prevention and intervention strategies and estimates of health care costs can be made. All of these depend on the accuracy and the coverage of the data. The mathematical studies can, in turn, point toward data needs and to insufficiencies in data quality."

4. "AIDS incidence data allow short-term projections (3 to 5 years). HIV prevalence is difficult to estimate. Statistical extrapolations have shown that AIDS appears to be increasing as the third power of the time. Dynamic model analysis has been used to exhibit the subgroup behavior accounting for this dependence. HIV prevalence has been obtained by back-calculation for special risk groups and the general population. However, error ranges are large for this kind of calculation. HIV incidence can only be measured in special groups."

5. "The quality of AIDS case data is good. Some selected risk group data are of special value. Improvements in the quality of HIV prevalence data require that the currently planned seroprevalence surveys, including those of the general population and the special family of surveys, proceed if they are found to be feasible."

6. "Progress toward long-range predictions for AIDS and HIV prevalence requires new biological and behavioral data and collaborative efforts between modelers, statisticians, epidemiologists, and behavioral, social and data management scientists."

7. "Collaboration has increased in the past year, and this has begun to increase the number of scientific results. In the past year there have been a number of meetings in which mathematical scientists, modelers, and statisticians have worked together with epidemiologists. Collaborative work has been funded by the National Institute for Drug Abuse (NIDA) on the advanced statistical analysis of AIDS among IV drug users at several universities."

8. "There is a good supply of high-quality, motivated mathematical researchers, applied mathematicians, and statisticians to conduct the modeling effort. There are currently about 20 people involved in all of the mathematics work, several of whom work in PHS. The work to be done will require a total of 60 people in the next several years."

9. "However, funding for the modeling effort is not adequate. Few of those now doing modeling are supported. Funding required to support those now working on modeling and the additional scientists required to carry out the needed research would amount to \$6 million by 1991."

10. "Access to AIDS and HIV data must be improved. As data are collected, they must be made available to researchers. A comprehensive directory of data sources is required to bring modelers, epidemiologists, and behavioral and social scientists to the data. Large-population, public-use databases must continue to be made more available in terms of the kinds of data released. Small-population datasets need to be made easier to access. A research environment that facilitates sharing of data needs to be promoted. However, all of these measures must be taken in a way that preserves confidentiality and maintains individual anonymity."

The ten conclusions of this 1988 Workshop are quite interesting. Despite the fact that the past contributions of epidemiological modeling cited in conclusion 1 are primarily conceptual, conclusion 3 emphasizes the use of mathematical modeling for forecasting. Although we are enthusiastic about the value of dynamic transmission modeling for conceptual understanding, we are more cautious about its value for forecasting future AIDS incidence and HIV prevalence and incidence. It may be difficult to do better forecasting than done by the extrapolation and back calculation techniques mentioned in conclusion 4 and described in more detail in the previous section. Dynamic transmission modeling can explore possible future incidences under different scenarios, but the predictions are likely to be similar to those obtained by short-term forecasting using simple direct methods such as extrapolation. We are disappointed that the Workshop findings put so much emphasis on the use of mathematical modeling to give forecasts. Most of the other findings seem valid since they address such issues as data and personnel availability, collaboration and funding.

Two more recent publications have concluded that mathematical models for HIV transmission and AIDS have great potential for forecasting, but that the data are not currently available for parameter estimation in these models.

In *AIDS: Sexual Behavior and Intravenous Drug Use* (Turner et al., 1989) published by the National Academy Press, it is stated on p. 74 that

"...promising mathematical models of the dynamics of the spread of HIV infections require data on a wide range of sexual behaviors; these data currently are not available. ... There are currently no reliable data on sexual contacts for the national population; there are also no such data for groups with elevated risks of transmitting or contracting HIV infections (e.g., men who have sex with men, IV drug users, heterosexuals with many sexual partners). Indeed, there is no reliable information on the size of the nonmonogamous heterosexual population. The lack of such data makes predictions about the future spread of AIDS extremely uncertain."

In *AIDS Forecasting*, published by the General Accounting Office (1989), four types of forecasting models are presented and analyzed. In their conclusions, they state on pp. 84-85 that

"...micro-simulation models were judged most comprehensive in the sense that they measured six components of the epidemic. These models explicitly consider HIV transmission and the development of AIDS and show how these contribute to past, current, and forecasted levels of the epidemic. In addition, projected future occurrences for individual components can be adjusted by the forecaster, allowing assessments of the effect of potential changes. But, unfortunately, the empirical basis of these models was extremely weak, due to lack of available information on specific components of the AIDS epidemic.

If knowledge about AIDS transmission were improved, the most comprehensive (micro-simulation) models would be deemed superior to the others in the sense that micro-simulation models make explicit the forces that contribute to the forecast."

We agree with the two quotations above; currently there are inadequate data to estimate the parameters in dynamic transmission simulation models well enough so that they will provide accurate forecasting. As better data become available, the forecasts of future HIV and AIDS incidences will improve, but it will be very difficult to estimate the accuracy of these forecasts. The social mixing, sexual interactions and needle-sharing behavior in our society are very complicated so that some simplifications are necessary in a model. The forecasts depend on the model used and it is difficult to decide how to choose a model with the appropriate amount of complexity. One guiding philosophy is to include the minimum structure to give reasonable fits to the data. Once the model is specified, it is then necessary to estimate the values of the parameters in the model. Even if the complete current and past sexual and needle-sharing behavior were known for each person in the United States, the problem would still remain of deciding on a reasonable model and then deciding how to summarize, group, categorize and condense the data in order to estimate parameter values in the model. Transmission dynamics models can be used to help identify trends, make general forecasts or estimate the uncertainty in the forecasts, but they are not more likely than other approaches to give accurate AIDS forecasts in the U.S. or elsewhere in the world. Both of the quotations above have focused on forecasting as the reason for AIDS modeling. We believe that the other benefits of epidemiological modeling described in the previous section are very important and are the aspects where mathematical modeling and computer simulations of HIV/AIDS can make the biggest contributions. In this monograph we emphasize conceptual understanding by the reconstruction of HIV/AIDS epidemics instead of forecasting future AIDS cases.