constant voltage stimulator as described (5). Total RNA (0.5 μg) was extracted [P. Chomcynski and N. Sacchi, *Anal. Biochem.* **162**, 156 (1987)], reverse transcribed by superscript ribonuclease H⁻ (Gibco BRL) with random primers, and amplified by PCR with primers corresponding to nucleotides 3476 through 3499 and 3611 through 3587 of the mouse L1 complementary DNA sequence of M. Moos *et al.* [*Nature* **334**, 701 (1989)] for 28 cycles at an annealing temperature of 57°C. The L1 fragment corresponded to the cytoplasmic domain and included the alternative splicing region characteristic of Schwann cell L1 [L. M. Moscoso and J. R. Sanes, *J. Comp. Neurol.* **352**, 321 (1995)]. This was done to detect possible Schwann cell contamination in these cultures (none was observed). The PCR products were analyzed on 6% polyacrylamide gels and quantified by imaging densitometry (Universal Imaging, West Chester, PA).

15. The competitive L1 mimic was constructed from a 225–base pair (bp) nonhomologous DNA fragment (Bam HI–Eco RI fragment of viral oncogene *v-erbB*) ligated to sequences complementary to the L1-specific primers used to amplify endogenous L1, according to the manufacturer's protocol (Clontech, Palo Alto, CA).
16. No difference was observed in total protein content, number of DRG neurons, mean diameter of soma, shape of the size-frequency histograms, or percentage of neurons staining for neurofilament (RT97) specific for large-type DRG neurons (about 75%) [S. N. Lawson, A. A. Harper, E. I. Harper, J. A. Garson, B. H. Anderton, *J. Comp. Neurol.* **228**, 263 (1984)] after 0, 0.1, or 1 Hz stimulation for 5 days.
17. We performed SDS–polyacrylamide gel electrophoresis as described (*10*) and quantified results by imaging densitometry, with polyclonal antibodies to L1 from mouse brain [F. G. Rathjen and M. Schachner, *EMBO J.* **3**, 1 (1984)] and NCAM antibody [E. Bock *et al.*, *ibid.* **4**, 2765 (1985)]. Schwann cells were not detected in these cultures, and the 230-kD band was distinct from the lower weight isoform of L1 expressed by Schwann cells (*6, 7*).
18. Axons were stained with antibody against neurofilament, and automated image analysis was used to quantify the number of axonal bundles intersecting a 315-μm linear transect in the side compartment, parallel to the partition and 100 to 200 μm from the barrier. Statistical comparisons by analysis of variance (ANOVA) were based on the mean number of fascicle intersections from 10 to 15 measurements in each dish (n = 26 dishes).
19. Schwann cells were prepared as described (*11*) and plated on DRG cultures (0.15×10^6 cells per dish) after 5 days of stimulation. Cultures were fixed with 4% paraformaldehyde 4 days later and stained with monoclonal antibody against the Schwann cell marker CNPase [T. J. Sprinkle *et al.*, *Brain Res.* **426**, 349 (1987)]. Quantitative comparisons were based on the mean number of Schwann cells per 0.13 mm^2 determined from at least 15 microscopic fields in each dish. Preincubation with antibody against L1 inhibited Schwann cell adhesion by >30% ($P < 0.009$; $n = 7$).
20. We are grateful to P. Nelson for continuous helpful discussions and to M. O'Donovan for reading an early draft of this report.

17 May 1995; accepted 18 September 1995

Scope of the AIDS Epidemic in the United States

Philip S. Rosenberg

Two-dimensional deconvolution techniques are used here to reconstruct age-specific human immunodeficiency virus (HIV) infection rates in the United States from surveillance data on acquired immunodeficiency syndrome (AIDS). This approach suggests that 630,000 to 897,000 adults and adolescents in the United States were living with HIV infection as of January 1993, including 107,000 to 150,000 women. The estimated incidence of HIV infection declined markedly over time among white males, especially those older than 30 years. In contrast, HIV incidence appears to have remained relatively constant among women and minorities. As of January 1993, prevalence was highest among young adults in their late twenties and thirties and among minorities. An estimated 3 percent of black men and 1 percent of black women in their thirties were living with HIV infection as of that date. If infection rates remain at these levels, HIV must be considered as endemic in the United States.

Through the use of deconvolution methods known as backcalculation, the national AIDS database compiled by the Centers for Disease Control and Prevention (CDC) and the distribution of the incubation period between infection with HIV and diagnosis with AIDS can be used to reconstruct the historical incidence of HIV infection that best accounts for the observed epidemic of AIDS cases (*1*). In recent years, the incidence of AIDS in the United States has slowed, with rates that are approaching a plateau (*2*). This trend likely reflects both reduced infection rates since the mid-1980s and widespread use of prophylactic therapies, which delay the onset of AIDS (*3*). Although the plateau in national AIDS cases would appear to be a favorable sign, optimism must be tempered by an appreciation of the dynamic nature of the epidemic. The HIV virus entered the United States during the late 1970s and spread rapidly during the early to mid-1980s. During this early period, there were large susceptible populations at risk over a broad range of ages. As the epidemic matured, one would expect that new entrants to at-risk populations—homosexual men, injection drug users, and high-risk heterosexuals—would tend to be young. Hence, it is plausible that the epidemic would stabilize with fewer infections occurring in recent years compared to the mid-1980s. For purposes of epidemic monitoring, therefore, a key question is whether the incidence rate among young adults has declined in comparison to the rate among persons of the same age in the past.

Although there are only limited data available to address this question by direct observation, extensions of the backcalculation approach (*4, 5*) allow one to estimate the age-specific incidence of infection from the age-specific incidence of AIDS. To avoid potential biases resulting from CDC's expansion of the AIDS case definition in 1993 (*2*), this approach was applied to cases among adults and adolescents who were diagnosed up to 1 January 1993 and who met the previous 1987 case definition (*6*). Incidence counts were adjusted for delays in reporting and estimates were inflated by 18% to reflect cases that will never be reported (*7*).

Although the overall rate of increase slowed after 1987, AIDS incidence trends differ according to birth cohort (Fig. 1A). Incidence among persons born before 1960 increased during the early to mid-1980s and then approached a plateau during 1991 to 1992, but AIDS incidence among persons born after 1960 was very low until 1986 and has increased steadily since then. This qualitative difference is apparent among men and women in each racial and ethnic group (Fig. 1B).

The infection rate function $\nu(s, a)$ specifies the number of infections per year at calendar time s among persons aged a years. Estimates of $\nu(s, a)$ were derived from age-specific AIDS incidence data on the basis of the incubation distribution. The fundamental convolution equation (*4, 5*) is given by

$$E(Y_{t,k}) = \int_{T_0}^{t} \nu(s,k-t+s) f(t-s \mid k-t+s,s)\,ds \quad (1)$$

where $E(Y_{t,k})$ is the expected number of cases occurring at calendar time t among persons aged k years at diagnosis, T_0 is the assumed start date of the epidemic, and $f(t \mid a, s)$ is the incubation period density function for persons aged a years at infection who were infected at time s. The incubation distribution varies by age to reflect that younger age is associated with slower progression (*8*) and by time to account for the increasing use of prophylactic therapies since 1987 (*1, 3–5*). Given observed AIDS incidence data $Y_{t,k}$ and an

National Cancer Institute, Rockville, MD 20852–4910, USA.

estimate of the family of incubation distributions $f(t|a, s)$ derived from cohort studies, $\nu(s, a)$ is estimated by maximum likelihood under the assumption that the observed counts followed a Poisson distribution (9).

The infection rate function was modeled with bivariate step functions that had relatively broad steps in the time dimension and narrow steps in the age dimension. This approach yielded a good fit to the observed data without making strongly parametric assumptions about the infection rates (9). The low hazard of AIDS in the first few years after infection makes it difficult to backcalculate recent HIV incidence (*1*). For this reason, a 6-year step was used to cover the 1987 to 1992 period. Use of a long last step implies that estimated incidence for the recent past reflects a broad period average, and no inferences can be made about trends within this period.

Conservative estimates of incidence were derived from a Weibull model of the incubation distribution with a 9-year median time-to-AIDS for individuals aged 30 years at infection (the "fast" model) (*10*). Larger estimates were derived from a "slow" model that incorporated a Weibull hazard with a change point to a linear slope (*11*) and with a median of 9.8 years to AIDS. A modest therapy effect on the incubation distribution was used to reflect that the efficacy of treatment appears to attenuate within about 1 year (*12*). Plausible ranges of estimates (*1*, 9) were calculated by subtracting two standard errors from point estimates derived from the fast incubation distribution and by adding two standard errors to point estimates derived from the slow incubation distribution. Given estimates of the numbers of incident HIV infections derived by backcalculation, surveillance data on AIDS mortality and census data on population size were used to estimate the age-specific incidence of HIV infection per 100,000 person-years at risk and to estimate the percentage of the population that was alive with HIV infection as of 1 January 1993 (*13*).

In white males, AIDS cases have been concentrated among men in their thirties since the onset of the epidemic, but in recent years, AIDS incidence has been relatively high among men in their late twenties (Fig. 2A). In each age group, the number of infections per year increased and then decreased over time, with peaks in the 1983 to 1986 period (Fig. 2B). Although all age groups showed a decline in the number of new infections over time, the reductions were substantially greater for men older than 30 years compared to younger men.

Estimates of the mean annual number of infections per 100,000 person-years at risk among white, black, and Hispanic men and

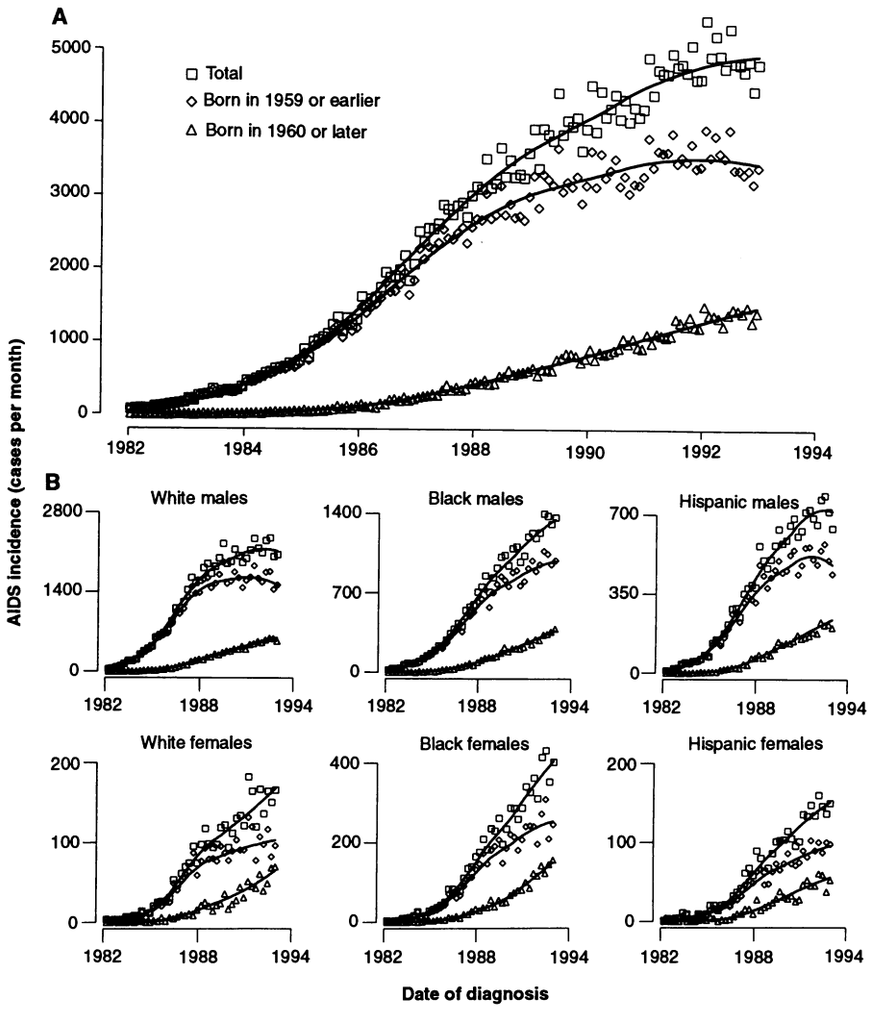

Fig. 1. (**A**) Monthly AIDS incidence counts for the total U.S. population (squares) and for individuals born in 1959 or earlier (diamonds) and in 1960 or later (triangles). (**B**) AIDS incidence for white, black, and Hispanic men and women. Locally weighted regression smoothing (*19*) with a bandwidth of 40% of the time axis was used to highlight trends in incidence (solid curves). The date axes mark 1 January of the years indicated.

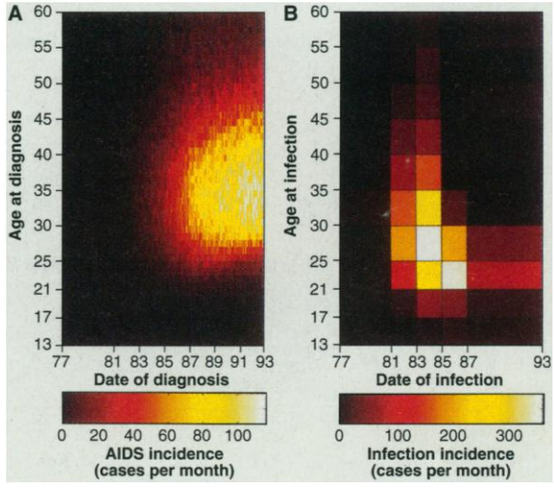

Fig. 2. (**A**) Monthly AIDS incidence counts among white males for January 1982 through December 1992 by single year of age (ages 13 to 59 are shown; ages 13 to 74 were used in the analyses). (**B**) Backcalculated number of infections per month by single year of age, calculated with the fast incubation distribution; results were qualitatively similar when the slow incubation distribution was used. Backcalculation models were fitted to the observed data after grouping the monthly periods into quarters and the single years of age into 17 age-at-diagnosis categories to reduce the computational burden. Appropriate expectations were calculated for the counts in these cells by integration of Eq. 1 (*9*). Rectangular regions in (B) correspond to areas of assumed constant infection intensity. Results were similar in a sensitivity analysis with 10 knots on the time axis based on cutpoints set to 1 January of calendar years 1977, 1981 through 1988, and 1993.

women for 1981 to 1986 and 1987 to 1992 show a rapid rise in infection incidence beginning in the late teens and early twenties and peak rates among individuals in their mid- to late twenties (Fig. 3). Only among white males was there a marked decline over time. Rates may also have declined among younger Hispanic men. Black females showed increasing infection rates.

Among white males, age-specific HIV prevalence as of 1 January 1993 peaked at about 1% among individuals in their early thirties (Fig. 4). Prevalence rates were high among minority men over a broad range of ages, 30 to 44 years. About 3% of black males and 1.5% of Hispanic males in this age range were living with HIV infection. Prevalence was about 1% for black females and 0.5% for Hispanic females in their late twenties and early thirties. Prevalence was lowest in white women.

Between the start of the epidemic and 1 January 1993, an estimated 857,000 to 1.1 million Americans had become infected with HIV but 227,000 had died of AIDS. The resulting plausible range for persons living with HIV infection as of 1 January 1993 was 630,000 to 897,000 (Table 1). The total includes an estimated 107,000 to 150,000 HIV-infected women, yielding an estimated male-to-female prevalence ratio for that date of about 4.6 to 1.

The backcalculation estimate of 630,000 to 897,000 prevalent infections in the entire population as of January 1993 is lower than the previous Public Health Service (PHS) estimate of around 1 million infected (*14*), a figure that has been widely used for planning purposes. The difference is less striking when one considers that the plausible range for the PHS estimate was 800,000 to 1.2 million (*14*), a range that overlaps with previous ranges derived by backcalculation (*1*, *15*) and the present study. Multiplying the incidence rates in Fig. 3 by the corresponding population sizes indicates that an average of 40,000 to 80,000 new infections occurred each year from 1987 to 1992, a range similar to previous estimates (*1*, *14*). However, this estimate provides no information about trends during the 6-year period and is subject to even greater uncertainty than the prevalence estimates.

The highest national prevalence rates were seen among minorities, and in particular among young black men. These estimates are derived from the large numbers of AIDS cases that have already been diagnosed. It is sobering to consider that 1 of every 50 black men in the United States aged 18 to 59 may be infected (Table 1), but a similar high rate in black men aged 18 to 59 of 1.9% was found in the Third National Household and Nutrition Examination Survey (*16*), a study that may have yielded conservative estimates because of nonresponse bias. The prevalence estimate derived here for women of 107,000 to 150,000 is also consistent with prevalence

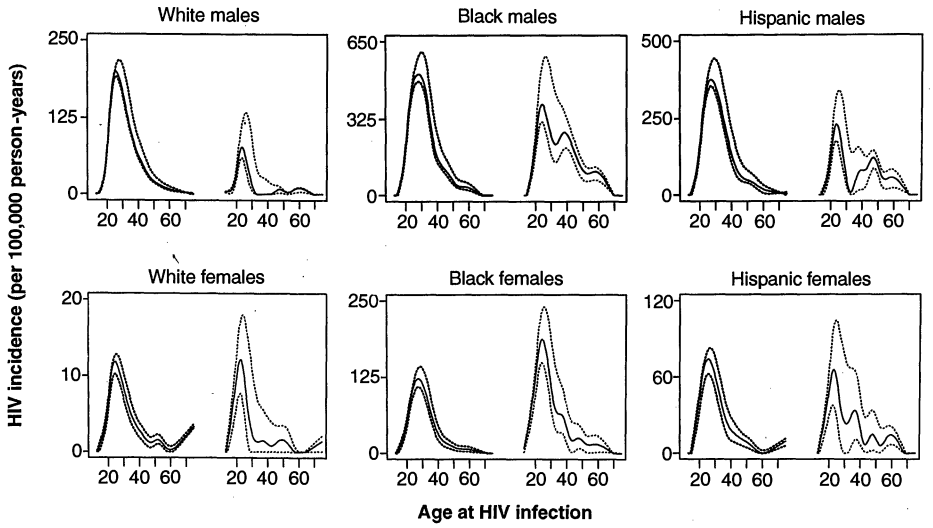

Fig. 3. Estimates of the number of infections per 100,000 person-years at risk by single year of age among white, black, and Hispanic men and women. Age-specific numbers of infections were calculated for each demographic group and for both models of the incubation distribution, as described in the legend to Fig. 2. For each model, the average infection rate per year was calculated for calendar periods 1981 through 1986 (left subplot) and 1987 through 1992 (right subplot), and a spline interpolant was drawn through the midpoints of the 11 age steps to highlight trends. Solid curves show point estimates based on the fast incubation distribution, and dotted curves show plausible ranges.

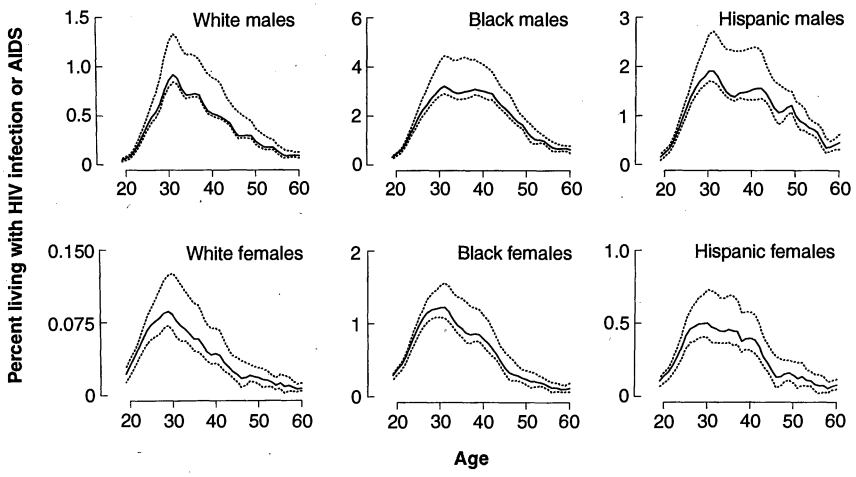

Fig. 4. Estimates of HIV prevalence as of 1 January 1993 by single year of age among white, black, and Hispanic men and women (*13*). Solid curves show point estimates based on the fast incubation distribution and dotted curves show plausible ranges. Summary prevalences for persons aged 18 to 59 years are shown in Table 1.

Table 1. Prevalence of HIV-1 infection in the United States as of 1 January 1993, estimated by backcalculation from AIDS incidence data (*7*) as described in the legends to Figs. 2 through 4. Point estimates are based on the fast model of the incubation distribution and may be conservative (*10*). Plausible ranges, given in parentheses, reflect some uncertainty about the incubation distribution.

Group	Alive with HIV-1 infection (in thousands)	HIV-1–positive, ages 18 to 59 (%)
	Males	
White	255 (248–370)	0.49 (0.48–0.70)
Black	184 (176–236)	2.29 (2.20–2.91)
Hispanic	97 (91–131)	1.44 (1.37–1.87)
Total	544 (523–747)	0.78 (0.77–1.04)
	Females	
White	25 (23–34)	0.05 (0.04–0.06)
Black	67 (63–82)	0.74 (0.69–0.90)
Hispanic	24 (21–32)	0.34 (0.31–0.45)
Total	117 (107–150)	0.16 (0.15–0.20)
Grand total	660 (630–897)	0.47 (0.46–0.62)

rates in the national Survey in Childbearing Women (*17*), which suggested that 130,000 women in the United States were infected as of 1992 (*18*).

It is important to recognize that backcalculated estimates are based on modeling rather than direct data and are very uncertain. The plausible ranges shown in Figs. 3 and 4 and in Table 1 account for a substantial amount, but not all, of the uncertainty about the incubation distribution and for random error from model fitting. These ranges do not account for uncertainty about the adjusted AIDS incidence counts or about the inflation factor used to account for cases that will never be reported, nor do they consider the choice of model of the infection rate function.

However, a central observation is that AIDS incidence increased much more rapidly in recent years among younger individuals born in 1960 or later than among older individuals (Fig. 1). Any proposed model must explain these qualitatively different epidemic curves. If infection rates remain at the levels indicated by these models, HIV must be considered an endemic infection affecting successive cohorts of young adults.

REFERENCES AND NOTES

1. R. Brookmeyer, *Science* **253**, 37 (1991).
2. Centers for Disease Control, *Morb. Mortal. Wkly. Rep.* **43** 826 (1994). This plateau was obscured after January 1993 by the expansion of case definition to include HIV-infected persons with severe CD4 depletion [Centers for Disease Control, *ibid.* **41** (no. RR-17) (1992)].
3. M. H. Gail, P. S. Rosenberg, J. J. Goedert, *J. AIDS* **3**, 296 (1990).
4. N. G. Becker and I. C. Marschner, *Biometrika* **80**, 165 (1993); A. Verdecchia *et al.*, *Int. J. Epidemiol.* **23**, 1027 (1994).
5. P. S. Rosenberg, *Stat. Med.* **13**, 1975 (1994).
6. Centers for Disease Control, *Morb. Mortal. Wkly. Rep.* **36** (suppl. 1S) (1987).
7. The analysis was based on all cases in the United States and its territories, including Puerto Rico, who were older than 13 years at diagnosis and who were reported to CDC as of March 1994. The AIDS incidence and mortality data for individuals who met the 1987 case definition (*6*) were adjusted for reporting delays with standard CDC adjustment weights. Incidence counts were inflated by 18%, an approximate figure, to account for unreported cases and for infected individuals who died before meeting the case definition. Adjusted totals of 336,000 incident cases of AIDS and 227,000 AIDS deaths as of 1 January 1993 were stratified by month of diagnosis, single year of age at diagnosis, gender, and race and ethnicity (non-Hispanic whites, non-Hispanic blacks, Hispanics, and others). Use of the 1987 definition maintains consistency between the AIDS data and available estimates of the incubation distribution.
8. M. E. Eyster *et al.*, *Ann. Intern. Med.* **107**,1 (1987); A. R. Moss *et al.*, *Br. Med. J.* **296**, 745 (1988); A. B. Mariotto *et al.*, *Am. J. Epidemiol.* **135**, 428 (1992); P. S. Rosenberg, J. J. Goedert, R. J. Biggar, *AIDS* **8**, 803 (1994).
9. Technical details of the fitting procedures are described elsewhere (*5*).
10. This fast model may yield conservative estimates because the natural history hazard function may level off around 7 years after infection [J. C. M. Hendriks *et al.*, *AIDS* **7**, 231 (1993)]. The Weibull hazard, in contrast, continues to increase. To reflect that progression rates appear to be slower with infection at younger ages, the age-specific hazard of progression for the fast model was assumed to increase (decrease) by the factor 1.042 (0.960) per year increase (decrease) in the age at infection compared to a person aged 30 years at infection, as derived in (*9*).
11. The estimated linear slope effect was 0.01 per year at 2.6 years after infection and the corresponding age effect was 1.037 (*9*). Both the fast and slow estimates of the incubation distribution were derived from a large cohort study of HIV-infected homosexual men [R. J. Biggar and the International Registry of Seroconverters, *AIDS* **4**, 1059 (1990)].
12. Therapy was assumed to reduce the hazard of AIDS among treated individuals by the factor 0.50 for persons infected for about 5 years, and then to wear off (*15*). The estimated cumulative proportion in treatment before AIDS was 40% for white males and 20% for other demographic groups. Therapy was assumed to have been introduced in April 1987 among white men and women and in April 1988 among other groups. These parameters are broadly consistent with clinical trial results [Concorde Coordinating Committee, *Lancet* **343**, 871 (1994)] and with estimates of the extent of therapy use in different groups [N. M. H. Graham *et. al.*, *J. AIDS* **4**, 267 (1991); P. S. Rosenberg *et al.*, *ibid.*, p. 392; W. Lang *et al.*, *ibid.* **6**, 191 (1993)].
13. To estimate age-specific prevalence at time T, note that individuals aged k years represent the surviving members of the cohort born at time $T - k$. Prevalence is estimated as the cumulative probability that a member of this cohort becomes infected as of T minus the cumulative probability that a member dies of AIDS as of T. These quantities can be estimated with actuarial formulas applied to census and AIDS mortality data and the results from the backcalculations.
14. Centers for Disease Control, *Morb. Mortal. Wkly. Rep.* **39**, 110 (1990).
15. P. S. Rosenberg, M. H. Gail, R. J. Carroll, *Stat. Med.* **11**, 1633 (1992).
16. G. M. McQuillan *et al.*, *J. AIDS* **7**, 1195 (1994).
17. M. Gwinn *et al.*, *J. Am. Med. Assoc.* **265**, 1704 (1991).
18. J. M. Karon, personal communication. This estimate is based on an analysis of unpublished CDC data.
19. W. S. Cleveland, *J. Am. Stat. Assoc.* **74**, 829 (1979).

20. I gratefully acknowledge M. Morgan for providing national AIDS surveillance data, T. Green for calculating the reporting delays, J. Smith of Information Management Systems for data management, and J. M. Karon and R. J. Biggar for many helpful discussions.

27 March 1995; accepted 14 September 1995