Subtype-specific Transmission Probabilities for Human Immunodeficiency Virus Type 1 among Injecting Drug Users in Bangkok, Thailand

Michael G. Hudgens,[1,2] Ira M. Longini, Jr.,[1] Suphak Vanichseni,[3] Dale J. Hu,[4] Dwip Kitayaporn,[5] Philip A. Mock,[6] M. Elizabeth Halloran,[1] Glen A. Satten,[4] Kachit Choopanya,[3] and Timothy D. Mastro[4,6]

The Bangkok (Thailand) Metropolitan Administration cohort of injecting drug users (IDUs) consisted of 1,209 IDUs initially seronegative for human immunodeficiency virus (HIV) who were followed from 1995 to 1998 at 15 Administration drug treatment clinics. At enrollment and approximately every 4 months thereafter, participants were assessed for HIV seropositivity. As of December 1998, there were 133 HIV type 1 seroconversions and approximately 2,300 person-years of follow-up. Of the 133 observed seroconversions, specimens from 126 persons were available for subtyping (27 subtype B, 99 subtype E). In this analysis, the authors assessed differences in subtype-specific transmission while controlling for important risk factors. The methodology used accounts for left truncation, interval censoring, and competing risks as well as for time-varying covariates such as each IDU's history of reported frequency of injection and of incarceration. Using plausible epidemiologic assumptions and controlling for behavioral risks, the authors found that a significantly higher transmission probability was associated with subtype E compared with subtype B in this population. Since many epidemiologic, virologic, and host factors can influence HIV transmission, it was difficult to conclude whether these differences in transmission probabilities were due to biologic properties associated with subtype. *Am J Epidemiol* 2002;155:159–68.

disease transmission; HIV; HIV infections; HIV-1; risk; risk factors; survival analysis

Human immunodeficiency virus type 1 (HIV-1) has been classified into several subtypes based on genetic variation (1). The explosive 1988 HIV-1 epidemic among injecting drug users (IDUs) in Bangkok, Thailand, was largely due to subtype B strains, while the larger heterosexual epidemic, which began expanding in 1989, was overwhelmingly due to subtype E strains (2–5). However, subtype E has accounted for an increasing proportion of new infections in Bangkok IDUs (6–8). Factors contributing to this increase in subtype E among IDUs are not well understood (9).

One question of interest is whether observed differences in subtype prevalence are due to immunologic or virologic rather than other factors (9, 10), such as demographic factors, socioeconomic factors, or geographic clustering. If indeed certain subtypes can be shown to be transmitted more efficiently than others, the implications on both scientific and public health policy could be important (9). The objective of this analysis was to assess differences in subtype-specific transmission in a prospective cohort study of IDUs potentially exposed to HIV-1 subtypes B and E. While much work has focused on sexual-contact transmission of human immunodeficiency virus (HIV) (e.g., refer to Gray et al. (11)), substantially less has focused on needle-sharing transmission (12–14). To our knowledge, there are no published estimates of the relative probabilities of subtype-specific needle-sharing transmission.

MATERIALS AND METHODS

Data

To better understand HIV transmission and to assess the feasibility of conducting a phase III HIV vaccine efficacy trial in an IDU population in Bangkok, the Bangkok Metropolitan Administration IDU cohort was established in 1995 (12). IDUs attending 15 Bangkok Metropolitan Administration drug treatment clinics were offered, with voluntary informed consent, enrollment in a prospective study. The study protocols were approved by the Ethical Review of Research Committee, Ministry of Public Health, Nonthaburi, Thailand, and an institutional review board at the Centers for Disease Control and Prevention, Atlanta,

Georgia. HIV-seronegative IDUs were enrolled during two time periods: May–November 1995 ($n = 499$) and May–December 1996 ($n = 710$). Participants received HIV prevention counseling and were assessed for HIV seroconversion at enrollment and approximately every 4 months thereafter. They were informed of their infection status and were offered enrollment and follow-up under a separate protocol if infected (15). At each visit, data were collected on demographics and risk behaviors since the last visit.

At the completion of the study, 1,124 participants had at least one follow-up visit. The overall HIV-1 incidence rate was 5.8 per 100 person-years, with 133 HIV seroconversions over approximately 2,300 person-years of follow-up (12). Of the people who seroconverted, 99 were infected with subtype E and 27 with subtype B. For the remaining 7 seroconverters, the virus was not typed, but we assumed that it would have been either subtype B or E. Enrollment times and seroconversion intervals for the persons who seroconverted during the study period are depicted in figure 1. The horizontal axis shows calendar time from the earliest enrollment date (May 30, 1995) to the study's completion at the end of 1998. In this figure, data for the 133 seroconverters were first sorted vertically by subtype and then by time of last seronegative test.

Covariates considered were gender, age (in years), jail history, history of injecting drugs in jail, frequency of needle sharing, frequency of needle injecting, and casual sex. Jail history was a dichotomous covariate indicating whether the study participant had been imprisoned since the last visit. Similarly, history of injecting drugs in jail indicated

whether, since the last visit, the IDU had injected drugs while incarcerated. The covariates needle-sharing frequency and needle-injecting frequency were recorded by using the following nine categories: 0) never/none, 1) less than once a month, 2) one to three times a month, 3) about once a week, 4) two to three times a week, 5) four to six times a week, 6) about once a day, 7) two to three times a day almost every day, and 8) four or more times a day almost every day. Casual sex was a binary covariate indicating whether the study participant had reported engaging in casual sex since the last visit. Table 1 gives the incidence of seroconversion (seroconversions per person-year) by subtype for each covariate level. For example, over intervals during which IDUs reported not being incarcerated, there were 92 seroconversions (16 of subtype B, 73 of subtype E, 3 of untyped virus) and 2,012.9 person-years of follow-up. In table 1 only, the continuous covariate age was partitioned into quartiles. A more comprehensive description of the cohort is given elsewhere (12).

Statistical methods

The Bangkok Metropolitan Administration cohort study gave rise to competing-risks failure-time data subject to interval censoring and left truncation. That is, we assumed that infection with one subtype precluded infection with a second subtype, such that Bangkok IDUs might be infected with either HIV-1 subtype B or E, but not both. Although dual infections with subtypes B and E have been reported (16), these infections do not appear to be common and were
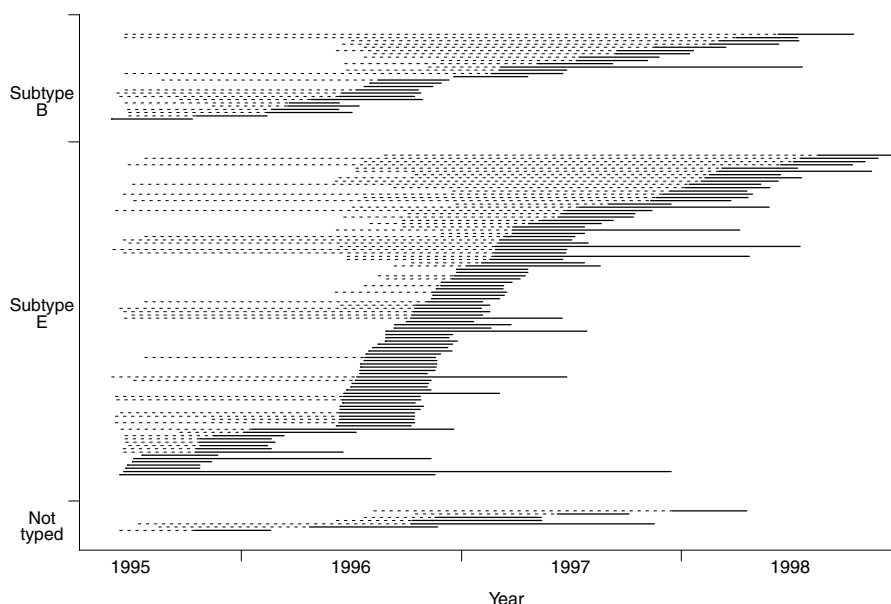
FIGURE 1. Graphic representation of those injecting drug users in the Bangkok (Thailand) Metropolitan Administration cohort study who seroconverted during the study period (1995–1998), by human immunodeficiency virus subtype. The thin dotted lines represent the time from enrollment to the last seronegative test. The bold lines represent the censoring interval (i.e., the time between the last seronegative and the first seropositive tests).

TABLE 1.   Incidence of seroconversion, by human immunodeficiency virus type 1 subtype, for each covariate level among injecting drug users in the Bangkok Metropolitan Administration cohort study, Thailand, 1995–1998

Covariate	Seroconversions (no.)				Person-years at risk	Incidence			
	Subtype B	Subtype E	NT*	Total		Subtype B	Subtype E	NT	Overall
Gender									
Male	25	87	7	119	2,179.6	0.011	0.040	0.003	0.055
Female	2	12	0	14	149.2	0.013	0.080	0.000	0.094
Jail									
No	16	73	3	92	2,012.9	0.008	0.036	0.001	0.046
Yes	11	26	4	41	315.9	0.035	0.082	0.013	0.130
Injecting while in jail									
No	24	86	4	114	2,257.3	0.011	0.038	0.002	0.051
Yes	3	13	3	19	71.5	0.042	0.182	0.042	0.266
Needle-sharing frequency†									
0	23	73	4	100	2,049.0	0.011	0.036	0.002	0.049
1	2	11	1	14	164.7	0.012	0.067	0.006	0.085
2	1	10	1	12	68.8	0.015	0.145	0.015	0.174
3	0	1	0	1	7.3	0.000	0.137	0.000	0.137
4	0	3	0	3	13.0	0.000	0.232	0.000	0.232
5	1	0	0	1	4.5	0.221	0.000	0.000	0.221
6	0	1	0	1	6.7	0.000	0.149	0.000	0.149
7	0	0	1	1	11.3	0.000	0.000	0.089	0.089
8	0	0	0	0	0.0	0.000	0.000	0.000	0.000
Needle-injecting frequency†									
0	0	1	0	1	114.8	0.000	0.009	0.000	0.009
1	5	11	2	18	450.3	0.011	0.024	0.004	0.040
2	0	10	0	10	305.1	0.000	0.033	0.000	0.033
3	1	3	0	4	138.5	0.007	0.022	0.000	0.029
4	0	10	2	12	202.6	0.000	0.049	0.010	0.059
5	1	2	0	3	55.7	0.018	0.036	0.000	0.054
6	6	11	1	18	285.7	0.021	0.039	0.004	0.063
7	12	36	2	50	597.0	0.020	0.060	0.003	0.084
8	2	15	0	17	179.2	0.011	0.084	0.000	0.095
Age (years)									
17–24	4	32	1	37	499.2	0.008	0.064	0.002	0.074
25–30	9	26	1	36	478.4	0.019	0.054	0.002	0.075
31–36	10	26	2	38	598.4	0.017	0.043	0.003	0.064
37–52	4	15	3	22	752.7	0.005	0.020	0.004	0.029
Casual sex									
No	27	93	5	125	2,148.1	0.013	0.043	0.002	0.058
Yes	0	6	2	8	180.6	0.000	0.033	0.011	0.044

   * Infecting virus not typed.
   † Frequency categories: 0, never/none; 1, less than once a month; 2, one to three times a month; 3, about once a week; 4, two to three times a week; 5, four to six times a week; 6, about once a day; 7, two to three times a day almost every day; and 8, four or more times a day almost every day.

not considered here. Because the time of seroconversion was known to be only between the last seronegative test and the first seropositive test, the event times were interval censored. Since HIV incidence may vary over time, the appropriate time scale for this analysis was calendar time starting at time $t = 0$, the earliest enrollment date (May 30, 1995). Because only those IDUs who were seronegative at the time of enrollment were included in the study, the seroconversion times were left truncated. In other words, IDUs who were seronegative at $t = 0$ may not have been included in the study because they had become seropositive by enrollment. In this section of the paper, we discuss statistical methods

ranging from nonparametric estimation to fully specified transmission probability models that allow quantification of differences between subtype B and E transmission. All methods are likelihood based, with individual likelihood contributions determined by enrollment time, seroconversion interval and subtype, and covariate information when appropriate. Persons for whom subtype was missing contributed to the marginal probability of seroconversion for subtype B or E.

Nonparametric exploratory methods.   Let T be the random variable for time of seroconversion, with cumulative distribution function (CDF) given by $F(t) = \Pr[T \leq t]$. This is

the probability of being infected (with any subtype) by time t for an IDU enrolled in the study at time 0. Let the cumulative incidence function for subtype B be $I_B(t) = \Pr[T \leq t,$ subtype B infection]. This is the probability of subtype B infection by time t when risk of infection from subtype E is present. Similarly, define $I_E(t)$ as the cumulative incidence function for subtype E so that, assuming that B and E are the only two subtypes circulating in this population, $I_B(t) + I_E(t) = F(t)$. The nonparametric maximum likelihood estimates of the CDF and the cumulative incidence functions for subtypes B and E are found by using methods derived previously (17–20).

Flexible parametric models. In this section, we introduce a flexible framework for modeling interval-censored and left-truncated competing-risks failure-time data that incorporates covariates. We begin with an additive hazards model, which assumes only that infections with subtypes B and E are distinct, mutually exclusive events. Note that additive hazards do not imply that the latent-variable survival times for subtypes B and E are independent (21). Allowing for dependent competing risks is particularly important, since it seems plausible that an increase in high-risk activities such as needle sharing would put a person at higher risk for infection from either subtype.

Let the baseline hazard functions for subtype B and E infections be $\lambda_B(t)$ and $\lambda_E(t)$, respectively, such that the overall baseline hazard function is given by $\lambda(t) = \lambda_B(t) + \lambda_E(t)$. We model the baseline hazard function at time t for subtype B infection by

$$\lambda_B(t) = \sum_{k=-3}^{K} S_k(t)\exp(\alpha_{Bk}),$$

where $S_k(t)$ are cubic-basis splines (22). The baseline hazard function for subtype E infection, $\lambda_E(t)$, is defined similarly. Without any covariates being considered, spline hazard functions can be used to estimate the cumulative distribution and incidence curves providing well-defined, smooth alternatives to nonparametric estimates.

Covariates were assumed to have a proportional effect on the subtype-specific hazard rates. That is, if $x(t)$ is some possibly time-dependent covariate, then the hazard function of subtype B infection at time t for a person with covariate $x(t)$ is

$$\lambda_B(t|x(t)) = \lambda_B(t)\exp\{\beta^B x(t)\}. \tag{1}$$

Covariates were similarly incorporated to modify the hazard function of subtype E infection. Note that the baseline hazard functions for subtypes B and E are not necessarily proportional. If we let the overall hazard function be $\lambda(t|x(t)) = \lambda_B(t|x(t)) + \lambda_E(t|x(t))$, it follows that

$$\lambda(t|x(t)) = \lambda(t)\exp\{\beta x(t)\} \tag{2}$$

only if a covariate has the same effect on each failure-type hazard function (i.e., $\beta^B = \beta^E = \beta$).

Transmission probability model. In contrast to the previous section, here the baseline hazard rates are given a specific structure based on the underlying infection process to allow estimation of the subtype-specific transmission probabilities. First, let the proportion of IDUs in Bangkok that are HIV infectious at time t be $\pi(t)$. Similarly, let the prevalence of subtypes B and E in the population of Bangkok IDUs at time t be $\pi_B(t)$ and $\pi_E(t)$, respectively, such that $\pi(t) = \pi_B(t) + \pi_E(t)$. Define a *contact* as a single act of sharing a syringe or needle with another IDU. Let the *sharing transmission probability* for subtype B be p_B, the probability of infection from a single sharing act with a subtype-B-infectious IDU. (The more general model in which the subtype-specific transmission probabilities are time dependent is not considered here.) Then, $p_B\pi_B(t)$ is the probability that a subtype B infection results from a single sharing act at time t with a randomly selected IDU of unspecified infectious status. Denote the contact (or sharing) rate at time t for a person by $c(t)$, given as number of contacts per day. Then, the subtype B infection hazard rate at time t for a person with contact rate $c(t)$ is

$$\lambda_B(t|c(t)) = p_B\pi_B(t)c(t). \tag{3}$$

Models similar to equation 3 have a rich history in infectious disease modeling (23). This model can be thought of as a proportional hazards model with a baseline hazard rate of $\lambda_B(t) = p_B\pi_B(t)$ and the log-transformed contact rate, that is, $x(t) = \log c(t)$, as the sole covariate having a coefficient equal to one. In model 3, the hazard rate of subtype B infection is zero when there are no contacts, that is, $c(t) = 0$. For subtype E, we define the sharing transmission probability, p_E, and the baseline hazard function, $\lambda_E(t) = p_E\pi_E(t)$, similarly such that $\lambda_E(t|c(t)) = \lambda_E(t)c(t)$. Then, the overall baseline hazard function can be written as $\lambda(t) = \lambda_B(t) + \lambda_E(t) = p(t)\pi(t)$, where the overall sharing transmission probability, $p(t) = p_B \times \pi_B(t)/\pi(t) + p_E \times \pi_E(t)/\pi(t)$, is the probability of infection from a single sharing act with an infectious IDU. The overall sharing transmission probability, $p(t)$, is an average of the subtype B and E transmission probabilities weighted by subtype proportions in the infectious population at time t. All covariates other than contact rate, $c(t)$, can be incorporated as in equations 1 and 2, for example, $\lambda_B(t|c(t), x(t)) = p_B\pi_B(t)c(t)\exp\{\beta^B x(t)\}$.

From the transmission probability model, our goal was to estimate the relative transmission probability, $R = p_E/p_B$, while controlling for covariate effects. To assess model fit, the coefficient on the log-transformed contact rate can be estimated instead of being held fixed at one. For example, we can allow the subtype B infection hazard to have the more general form

$$\lambda_B(t|c(t),x(t)) = p_B\pi_B(t)c(t)^{\gamma}\exp\{\beta^B x(t)\}. \tag{4}$$

Estimates of γ that differ significantly from one may be attributable to heterogeneity of IDUs in susceptibility or infectiousness, measurement error, or unmeasured confounding covariates (24).

Subtype distributions. Fitting the transmission probability model requires knowledge of the prevalence of different subtypes in the IDU population from which the study sample was drawn. On the basis of previous estimates (25), we assumed that overall HIV-1 prevalence was constant and 30

percent, that is, $\pi \equiv \pi(t) = 0.3$. However, the distribution of subtypes in circulation may change over time. Table 2 provides estimates of the proportion of subtype E from Bangkok Metropolitan Administration screening data over different time periods (25, 26). To our knowledge, no other estimates of the proportion of subtype E in the population of Bangkok IDUs during this time period are available. On the basis of the data in table 2, we assumed that the proportion of subtype E followed a piecewise linear form passing through each of the estimated proportions (figure 2). To account for uncertainty in the proportion of subtype E over time, a sensitivity analysis was performed by investigating the effect of a vertical shift Δ of the proportion of subtype E on the relative and subtype-specific estimates of transmission probabilities. Figure 2 depicts the proportion of subtype E over time for $\Delta = -0.10, -0.05, 0.05, 0.10$. The two

extremes (–0.10 and 0.10) give rise to estimates of the proportion of subtype E well beyond the 95 percent confidence intervals at each time point in the screening phase. Finally, note that these screening phases did not involve random samples of IDUs in Bangkok, since only those IDUs not known to be HIV positive were screened. The effect of this potentially biased sampling is examined further in the Discussion section of this paper.

Reported contact rate adjustment.   Thus far, the transmission probability model assumes that the only source of HIV exposure for the IDUs under study is needle sharing. Unfortunately, under this assumption, the reported sharing rate appears to be underreported; 100 of the 133 seroconversions occurred over intervals during which the IDUs reported no sharing (table 1). That $c(t)$ was mismeasured prohibits using the sharing rate directly in the model.

TABLE 2.   Proportion of injecting drug users in Bangkok, Thailand, with human immunodeficiency virus type 1 subtype E observed in the 1995–1996 screening phase of the Bangkok Metropolitan Administration cohort study and from additional screening in 1999 and 2000

Screening period	No. with subtype E	Total no.	Proportion with subtype E	95% CI*
May 1995–November 1995	68	194	0.35	0.28, 0.42
May 1996–December 1996	344	603	0.57	0.53, 0.61
March 1999–December 1999	419	569	0.74	0.70, 0.77
January 2000–August 2000	745	963	0.77	0.75, 0.80

* CI, confidence interval.
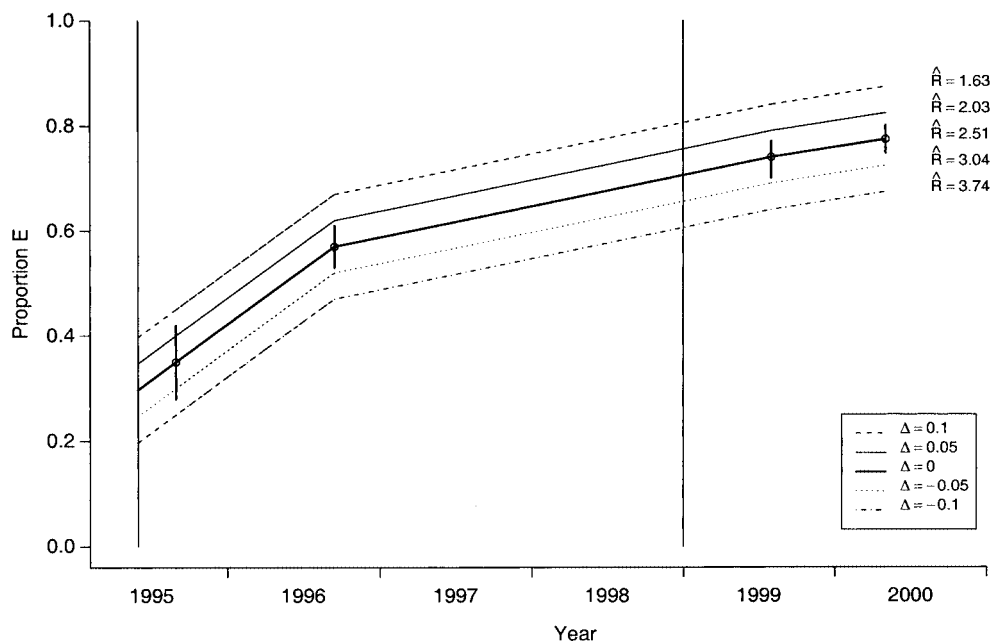
FIGURE 2.   Human immunodeficiency virus subtype distribution scenarios considered on the basis of the 1995–1996 screening phase of the Bangkok (Thailand) Metropolitan Administration cohort study of injecting drug users and subsequent screening phases in 1999 and 2000. The two vertical lines represent the beginning and end of the cohort study. The bold line is the piecewise linear model for the proportion of subtype E based on the screening-phase data ($\Delta = 0$). The bold vertical bars are the 95% confidence intervals for each screening-phase proportion. The thinner lines are the proportions of subtype E for different values of Δ in sensitivity analysis. Corresponding estimates of the relative transmission probability, *R*, from the unadjusted model in table 4 are also given.

However, despite being mismeasured, the reported sharing rate did contain some valid exposure information. For example, those IDUs who reported no sharing were at less risk than those IDUs who reported at least some sharing. On the other hand, the reported injecting rate appears to be a better measure of exposure to infection than the reported sharing rate since the rate of seroconversion generally increased with the former but did not increase with the latter, and all but one of the seroconversions occurred over intervals during which the reported injecting rate was not zero. Therefore, we made an adjustment that estimated the true sharing rate by using the reported individual-level injecting rates and a population-averaged sharing proportion.

Let the injecting rate at time t be denoted by $h(t)$ and assume that the sharing rate is some proportion of the injecting rate such that $c(t) = \theta h(t)$, where θ is the average proportion of injecting acts that entail sharing. Without other information about the true contact rate, θ, p_E and p_B are not identifiable since $c(t)$ always appears in the likelihood function as a product with either p_E or p_B. While the error in the reported sharing rate prohibits direct use in the model, we can use $c(t)$ to estimate θ by averaging the reported sharing rate to the injecting rate ratio, that is, $c(t)/h(t)$, over all visits for all persons. The true sharing rate is then estimated by $\hat{\theta} h(t)$. Refer to Hudgens et al. (13) for details. This reported contact rate adjustment should not bias $\hat{R}$ unless the degree of mismeasurement of the sharing rates is different across subtypes.

RESULTS

Nonparametric estimates of the CDF and the cumulative incidence curves are given in figure 3. The nonparametric

maximum likelihood estimate of the CDF can be inconsistent and can severely overestimate the cumulative probability of failure in the presence of left truncation because few persons are at risk when the study begins (27). Because only six persons were enrolled on the first study day, the nonparametric maximum likelihood estimate of the CDF increased 0.17 on day 0 (i.e., May 30, 1995), yielding an overestimate of the probability of infection from either subtype B or E. To compensate, the nonparametric estimate of the CDF shown in figure 3 is conditional on escaping infection during day 0. This is also true for the nonparametric estimates of $I_B(t)$ and $I_E(t)$. Unconditional B-spline estimates of the CDF and cumulative incidence functions are also given in figure 3. The agreement between the nonparametric and the B-spline estimators confirms that the increase in the CDF at day 0 is attributable to a small risk set and not extraordinarily high transmission rates at the beginning of the study.

B-spline estimates of the subtype-specific hazard function curves in which covariates were ignored are given by the bold lines in figure 4. Next, the data were fit to the B-spline regression model by using the covariates gender, age, incarceration, and injecting drugs in jail. To control for possible exposure to infection in a fashion similar to that in the transmission probability model, the injecting rate was also included in the model by using a log transformation with the regression coefficient fixed at one. Estimates of covariate effects and the corresponding p values from this B-spline regression model are given in the third and fifth columns of table 3, respectively. As shown previously (12, 13), being female, having been incarcerated, and having injected drugs while in jail were each associated with an increased risk of
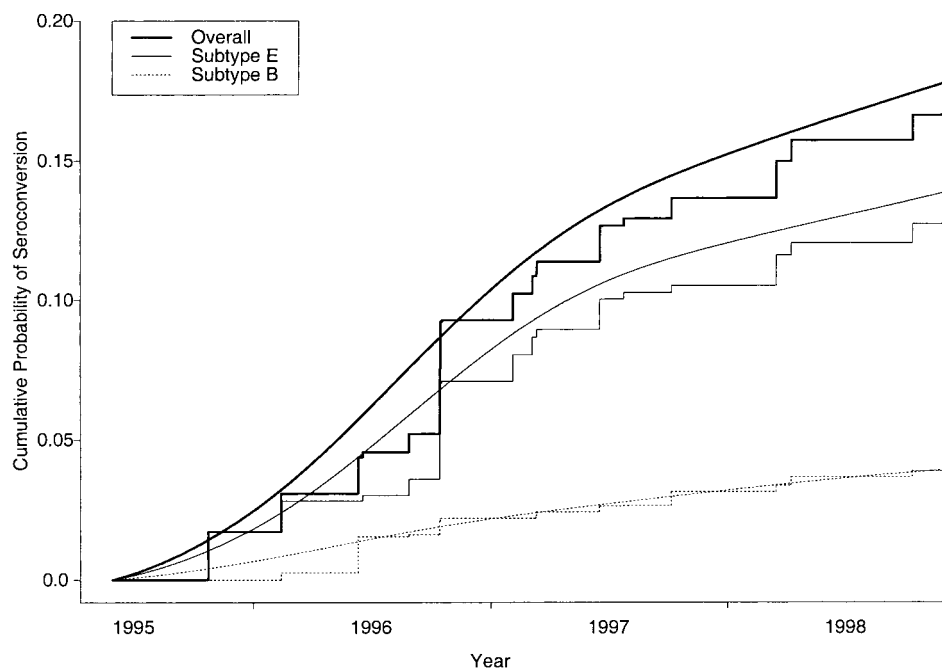
FIGURE 3. Nonparametric and B-spline estimates of the cumulative distribution and subtype-specific cumulative incidence functions for time to seroconversion for injecting drug users in the Bangkok (Thailand) Metropolitan Administration cohort during the study period 1995–1998.
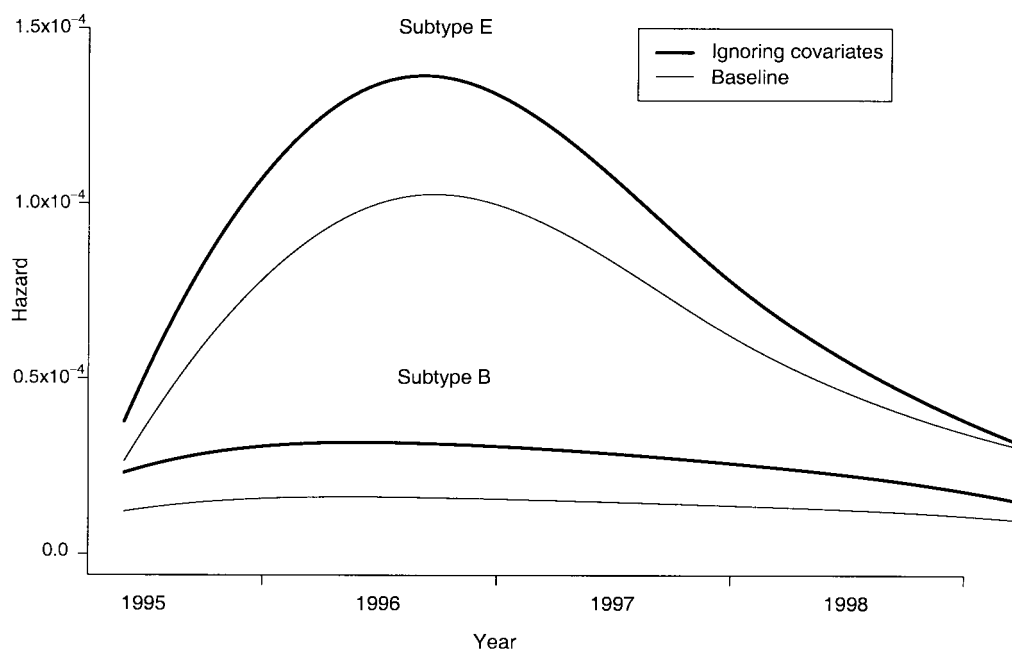
FIGURE 4. Estimate of human immunodeficiency virus subtype-specific hazard functions for injecting drug users in the Bangkok (Thailand) Metropolitan Administration cohort study, 1995–1998. The bold line represents hazard functions for which covariates were ignored. The thin line is the baseline hazard function from the B-spline regression model.

infection. The effect of gender on the subtype E hazard function was marginally significant (p value = 0.1), while there was no significant difference (p value > 0.8) between males and females for infection from subtype B. The effects of incarceration and injecting drugs while in jail did not differ significantly between subtypes. Casual sex was not found to be associated with risk of seroconversion in this population (p value = 0.49). The estimated covariate effects here are consistent with those shown in table 1, previous analyses of this cohort (12, 13), and the subtype transmission probability model discussed below. The resulting B-spline estimate of the subtype-specific baseline hazard functions associated with this model are given by the thin lines in figure 4. Even after covariates were accounted for, the peak in the estimate of the subtype E baseline hazard function persisted. Thus, this peak could not be fully explained by the time-varying covariates incarceration status, injecting drugs while in jail, or injecting frequency.

Based on the data, $\hat{\theta}$ = 0.04; that is, on average about 4 percent of injecting acts entail needle sharing. Conditional on $\hat{\theta}$ and the piecewise linear subtype proportion distribution model in figure 2, we fit the transmission probability model to the data. The estimated covariate effects, 95 percent confidence intervals, and p values given in the last three columns of table 3 are in agreement with the corresponding estimates from the B-spline regression model. Casual sex was not a significant predictor of risk for seroconversion (p value = 0.58). The reported sharing rate was not significant either (p value = 0.24), suggesting that exposure information is better captured by using injecting frequency and estimated sharing proportion. Estimates of the subtype transmission probabilities and the relative transmission probability are given in the first row of table 4, indicating that p_E is about 2.5 times greater than p_B. Estimates in the four rows that follow are part of the sensitivity analysis; the corresponding relative transmission probability estimates

TABLE 3. Covariate effect estimates for the B-spline and transmission probability models for injecting drug users enrolled in the Bangkok Metropolitan Administration cohort study, Thailand, 1995–1998

Covariate	Parameter	B-spline regression model			Transmission probability model		
		Estimate	95% CI*	p value	Estimate	95% CI	p value
Female gender	$\exp\{\beta^E_{Gender}\}$	1.60	0.89, 2.87	0.10	1.68	0.87, 2.98	0.12
Jail	$\exp\{\beta_{Jail}\}$	1.62	1.05, 2.46	0.04	1.66	1.01, 2.63	0.04
Injecting while in jail	$\exp\{\beta_{Jail + Inject}\}$	3.79	2.31, 5.91	<0.0001	4.47	2.63, 7.19	<0.0001
Age	$\exp\{\beta^E_{Age}\}$	0.99	0.96, 1.02	0.35	0.98	0.96, 1.02	0.23

* CI, confidence interval.

TABLE 4. Transmission probability estimates for different subtype distributions of human immunodeficiency virus type 1 among injecting drug users in the Bangkok Metropolitan Administration cohort study, Thailand, 1995–1998

Model	Δ*	p_B† × 10²		p_E‡ × 10²		R§	
		Estimate	95% CI¶	Estimate	95% CI	Estimate	95% CI
Unadjusted#	0.00	0.45	0.29, 0.65	1.12	0.87, 1.41	2.51	1.65, 3.94
	0.10	0.59	0.38, 0.85	0.96	0.74, 1.20	1.63	1.07, 2.58
	0.05	0.50	0.33, 0.73	1.03	0.80, 1.29	2.03	1.33, 3.21
	−0.05	0.40	0.26, 0.58	1.22	0.95, 1.54	3.04	2.01, 4.84
	−0.10	0.36	0.23, 0.52	1.35	1.05, 1.70	3.74	2.44, 5.93
Adjusted**	0.00	0.63	0.41, 0.92	1.57	1.21, 1.98	2.48	1.63, 3.88
	0.10	0.82	0.54, 1.21	1.34	1.03, 1.69	1.64	1.05, 2.54
	0.05	0.72	0.47, 1.04	1.44	1.11, 1.83	2.00	1.31, 3.16
	−0.05	0.57	0.37, 0.82	1.71	1.33, 2.16	3.02	1.97, 4.77
	−0.10	0.51	0.33, 0.75	1.89	1.46, 2.40	3.70	2.42, 5.81

* Δ, sensitivity analysis vertical shift in proportion E.
† p_B, subtype B transmission probability.
‡ p_E, subtype E transmission probability.
§ R, relative transmission probability p_E/p_B.
¶ CI, confidence interval.
\# Log contact rate coefficient γ fixed at one.
** γ estimated.

are also given in figure 2. Under all scenarios considered in the sensitivity analysis, the relative transmission probability was significantly greater than one, affirming that $p_E > p_B$.

The last five rows of table 4 give estimates of the transmission probabilities when the contact rate exponent (γ from equation 4) is estimated instead of being held fixed at one. For the piecewise linear subtype proportion model (Δ = 0), $\hat{\gamma} = 0.16$, which is significantly less than one, with a 95 percent confidence interval of 0.06, 0.25. Similar estimates of γ (not shown) were obtained under the sensitivity analysis scenarios. It was difficult to determine the source of this lack of fit, but certainly the mismeasured sharing rates offer a plausible explanation. While the subtype-specific transmission probability estimates are greater than the unadjusted model (γ fixed at one), the relative transmission probability estimates are nearly identical. Similarly, covariate effects (not shown) did not differ substantially from the data in table 3 for either the unadjusted or the adjusted models under the different subtype proportion scenarios considered.

DISCUSSION

Results from the various methods used in this analysis were all consistent with a higher transmission rate associated with subtype E infection, leading to the conclusion that the transmission probability per needle-sharing contact is likely greater for subtype E than for subtype B (i.e., $p_E > p_B$). It was difficult to conclude whether the difference in transmission probabilities was due to biologic properties associated with subtype, since epidemiologic, immunologic, and virologic factors can influence HIV transmission (28–30). For this analysis, the validity of the subtype-specific and relative transmission probability estimates relied on knowledge of the circulating subtype proportions. Fortunately, external estimates of the subtype prevalences were available. Use of the screening results to estimate the proportion

of subtype E in the infected population could have introduced a bias, since only those not known to be HIV positive were screened. However, making plausible assumptions, we can quantify the direction of this bias. If we assume that those persons found to be HIV positive during the screening phase represent more recently infected IDUs and that more-recent infections are more likely to be of subtype E, then it follows that the proportion of IDUs infected with subtype E is greater among those not known to be infected (i.e., the screening phase participants) than among the overall infected population. This conclusion implies that the estimate of subtype E prevalence from the screening phases is an upper bound for the true proportion of subtype E in the population. The resulting estimate of the relative transmission probability, $\hat{R} = 2.5$, would thus be an underestimate.

Another assumption of the transmission probability model is that the only source of HIV-1 exposure in the Bangkok Metropolitan Administration cohort was needle sharing. Both this and previous analysis of the same cohort data (12) suggest that transmission through needle sharing accounts for the large majority of infections with both subtypes B and E and that sexual behavior is not associated with increased risk of HIV-1 seroconversion. In fact, an increase in the frequency of casual sex was found to be inversely associated with the risk of HIV-1 seroconversion. Even if a few volunteers were infected by sexual transmission, this should not bias the estimate of R unless these seroconversions were unevenly distributed between subtypes.

The dominance of subtype E in this study may be due to a lack of independence among the observations. For example, it is well known that needle-sharing networks exist. Many of the IDUs in a network could have been infected with the same subtype at roughly the same time. Since there were 133 seroconverters in this study, a network of 20 or more could have a dramatic effect on the results. Further analysis should be conducted to assess the amount of clus-

tering, either by clinic, geographically, or by phylogenetic relation, of the infecting viral strains.

Other evidence suggests possible inherent biologic differences between subtypes B and E that may influence transmissibility. First, a cross-sectional study in Thailand suggested that subtype E was more sexually transmissible than subtype B (31), although this result may have been confounded by the risk category of the index case (9). Second, in another analysis of IDUs who seroconverted in the current cohort study, those infected with subtype E had higher mean plasma viral loads than those with subtype B, especially in the early period following seroconversion (15). Higher viral loads are thought to contribute to higher rates of transmission (32–35). The magnitude of the observed difference in viral load depends on the distribution of the seroconversion times for subtypes E and B. It is possible that over the time period of this study, Bangkok IDUs with subtype E might have been more recently infected, on average, than those with subtype B and hence more likely to be in their primary (peak) viremic phase (36–38), which would then contribute to higher transmission probabilities. Third, a recent study by Jeeninga et al. (39) suggested that subtype E may replicate more efficiently than subtype B in certain circumstances. These studies, coupled with ecologic observations in the Thailand IDU population showing changes in the distribution of subtypes over time (6–8), lend support to a higher needle-sharing transmission probability associated with subtype E as compared with subtype B.

ACKNOWLEDGMENTS

REFERENCES

1. Hu DJ, Dondero TJ, Rayfield MA, et al. The emerging genetic diversity of HIV: the importance of global surveillance for diagnostics, research, and prevention. JAMA 1996;275:210–16.
2. Weniger BG, Limpakarnjanarat K, Ungchusak K, et al. The epidemiology of HIV infection and AIDS in Thailand. AIDS 1991;5:S71–S85.
3. Ou CY, Takebe Y, Weniger BG, et al. Independent introduction of two major HIV-1 genotypes into distinct high risk populations in Thailand. Lancet 1993;341:1171–4.
4. Kalish ML, Baldwin A, Raktham S, et al. The evolving molecular epidemiology of HIV-1 envelope subtypes in injecting drug users in Bangkok, Thailand: implications for HIV vaccine trials. AIDS 1995;9:851–7.
5. Limpakarnjanarat K, Ungchusak K, Mastro TD, et al. The epidemiologic evolution of HIV-1 subtypes B and E among heterosexuals and injecting drug users in Thailand, 1992–1997. AIDS 1998;12:1108–9.
6. Wasi C, Herring B, Raktham S, et al. Determination of HIV-1 subtypes in injecting drug users in Bangkok, Thailand, using peptide-binding enzyme immunoassay and heteroduplex mobility assay: evidence of increasing infection with HIV-1 subtype E. AIDS 1995;9:843–9.
7. Subbarao S, Limpakarnjanarat K, Mastro TD, et al. HIV-1 in Thailand, 1994–1995: persistence of two subtypes with low genetic diversity. AIDS Res Hum Retroviruses 1998;14:319–27.
8. Subbarao S, Vanichseni S, Hu DJ, et al. Genetic characterization of HIV-1 subtype E and B strains from a prospective cohort of injecting drug users in Bangkok, Thailand. AIDS Res Hum Retroviruses 2000;16:699–707.
9. Mastro TD, Kunanusont C, Dondero TJ, et al. Why do HIV-1 subtypes segregate among persons with different risk behaviors in South Africa and Thailand? AIDS 1997;11:113–16.
10. Hu DJ, Buvé A, Baggs J, et al. What role does HIV-1 subtype play in transmission and pathogenesis? An epidemiological perspective. AIDS 1999;13:873–81.
11. Gray RH, Wawer MJ, Brookmeyer R, et al. Probability of HIV-1 transmission per coital act in monogamous, heterosexual HIV-1-discordant couples in Rakai, Uganda. Lancet 2001;357:1149–53.
12. Vanichseni S, Kitayaporn D, Mastro TD, et al. Continued high HIV-1 incidence in a vaccine trial preparatory cohort of injecting drug users in Bangkok, Thailand. AIDS 2001;15:397–405.
13. Hudgens MG, Longini IM, Choopanya K, et al. Estimating the HIV-1 transmission probability in injecting drug users in Thailand. Appl Stat 2001;50:1–14.
14. Kaplan EH, Heimer R. A model-based estimate of HIV infectivity via needle sharing. J Acquir Immune Defic Syndr 1992;5:1116–18.
15. Hu DJ, Vanichseni S, Mastro TD, et al. Viral load differences in early infection with two HIV-1 subtypes. AIDS 2001;15:683–91.
16. Artenstein AW, Van Cott TC, Mascola JR, et al. Dual infection with human immunodeficiency virus type I of distinct envelope subtypes in humans. J Infect Dis 1995;171:805–10.
17. Peto R. Empirical survival curves for interval censored data. Appl Stat 1973;22:86–91.
18. Turnbull B. The empirical distribution function with arbitrarily grouped, censored and truncated data. J R Stat Soc (B) 1976;52:290–5.
19. Frydman H. A note on nonparametric estimation of the distribution function from interval-censored and truncated observations. J R Stat Soc (B) 1994;56:71–4.
20. Hudgens MG, Satten GA, Longini IM. Nonparametric maximum likelihood estimation for competing risks survival data subject to interval censoring and truncation. Biometrics 2001;57:74–80.
21. Tsiatis AA. Competing risks. In: Armitage P, Colton T, eds. Encyclopedia of biostatistics. New York, NY: John Wiley & Sons, 1998:824–34.
22. Rosenberg PS. Hazard function estimation using B-splines. Biometrics 1995;51:874–87.
23. Halloran ME. Concepts of infectious disease epidemiology. In: Rothman KJ, Greenland S, eds. Modern epidemiology. 2nd ed. Philadelphia, PA: Lippincott Raven, 1998.
24. Shiboski SC, Jewell NP. Statistical analysis of the time dependence of HIV infectivity based on partner study data. J Am Stat Assoc 1992;87:360–72.
25. Kitayaporn D, Vanichseni S, Mastro TD, et al. Infection with HIV-1 subtypes B and E in injecting drug users screened for enrollment into a prospective cohort in Bangkok, Thailand. J Acquir Immune Defic Syndr Hum Retrovirol 1998;19:289–95.
26. Choopanya K, Tappero J, Pitisuttithum P, et al. Completion of enrollment and ongoing follow-up of injecting drug users (IDUs) in the AIDSVAX-B/E vaccine efficacy trial in Bangkok, Thailand. Presented at the 6th International Congress on AIDS in Asia and the Pacific, Melbourne, Australia, October 5–10, 2001.
27. Pan W, Chappell R. A note on inconsistency of NPMLE of the distribution function for left truncated and case I interval censored data. Lifetime Data Anal 1999;5:281–91.
28. Laga M, Nzila N, Goeman J. The interrelationship of sexually

transmitted diseases and HIV infection: implications for the control of both epidemics in Africa. AIDS 1991;5:S55–S63.

29. Mastro TD, de Vincenzi I. Probabilities of sexual HIV-1 transmission. AIDS 1996;10:S75–S82.

30. Morris M, Kretzschmar M. Concurrent partnerships and the spread of HIV. AIDS 1997;11:641–8.

31. Kunanusont C, Foy HM, Kreiss JK, et al. HIV-1 subtypes and male-to-female transmission in Thailand. Lancet 1995;345: 1078–83.

32. Mock PA, Shaffer N, Bhadrakom C, et al. Maternal viral load and timing of mother-to-child HIV transmission, Bangkok, Thailand. AIDS 1999;13:407–14.

33. Shaffer N, Roongpisuthipong A, Siriwasin W, et al. Maternal virus load and perinatal human immunodeficiency virus type 1 subtype E transmission, Thailand. J Infect Dis 1999;179:590–9.

34. Quinn TC, Wawer MJ, Sewankambo N, et al. Viral load and heterosexual transmission of human immunodeficiency virus type 1. N Engl J Med 2000;342:921–9.

35. De Cock KM, Adjorlolo-Johnson G, Ekpini E, et al. Epidemiology and transmission of HIV2: why there is no HIV-2 pandemic. JAMA 1993;270:2083–6.

36. Jacquez JA, Koopman JS, Simon CP, et al. Role of the primary infection in epidemics of HIV infection in gay cohorts. J Acquir Immune Defic Syndr 1994;7:1169–84.

37. Satten GA, Mastro TD, Longini IM. Modelling the female-to-male per-act HIV transmission probability in an emerging epidemic in Asia. Stat Med 1994;13:2097–106.

38. Mastro TD, Satten GA, Nopkesorn T, et al. Probability of female-to-male transmission of HIV-1 in Thailand. Lancet 1994;343:204–7.

39. Jeeninga RE, Hoogenkamp M, Armand-Ugon M, et al. Functional differences between the long terminal repeat transcriptional promoters of human immunodeficiency virus type I subtypes A through G. J Virol 2000;74:3740–51.