

HIV-1 Superinfection Surveillance in an Acute Infection Cohort using *pol* Sequences from Resistance Genotyping: 1996-2009

Larry Bragg¹, Jeff McConnell¹, Teri Liegler², Peter Bacchetti², Gerald Spotts², Frederick Hecht², Robert Grant^{1,2}

¹Gladstone Institute of Virology and Immunology, ²University of California, San Francisco.

Background

Sequential expression of dual infections (SEDI) has been reported more often in recently infected persons and is usually attributed to HIV-1 superinfection. The frequency, risk factors, and consequences of superinfection are still unknown. Few cases have been linked to a source partner and so underlying primary dual or multiple infection cannot fully be ruled out. A routine screening for apparent superinfection incidence was incorporated into the OPTIONS Study in 2004, a cohort of individuals with recent HIV-1 infection. Annual surveillance was started in 2008, both retrospectively and prospectively, with the aim of systematically investigating superinfection or multiple infection incidence and seroadaptive behaviors that may contribute to the risk of superinfection in recently infected persons.

Methods

Acute and recent HIV-1 infection was diagnosed in the San Francisco OPTIONS Study using estimation by serological evolution (RNA+AB-, detuned ELISA, western blot evolution, seroconversion in last 6 months). Population sequencing of the *pol* gene was performed (TRUGENE, Bayer) at enrollment and four subsequent query intervals (2004, 2005, 2008, & 2009) for superinfection surveillance if viral load > 100 copies/ml. Phylogenetic analysis of baseline and follow-up timepoints was conducted using bootstrapped neighbor joining trees (Clustal X v2.06). Individual's sequences were considered related if their sequences were located on the same branch with a genetic distance < 6%. Person-years of observation were assigned to year of infection by first estimating infection dates (EID, AIEDRP algorithm). The person-years of observation attributed to the first year of infection were calculated by subtracting the baseline genotype date (usually

Methods (cont.)

screening or enrollment) from a date 365 days from EID to measure the duration of observation. Observation was attributed to the second and subsequent years of infection by measuring how many individuals were followed between 366 and 731 days (2nd year), 732 and 1,097 (3rd year), and so on. Individuals lost to follow-up or going on treatment were assigned partial years based upon their last follow-up genotype. The off treatment observation was calculated by subtracting any time on treatment from the person-years observed by year of infection. Observed SEDI events were assigned by year of infection using the date when the second or divergent virus was first observed by population genotype.

Statistical

We analyzed occurrence of SEDI using time-to-event (survival analysis) methods, employing the simplest parametric model, exponential, because of the small number of SEDI events. To evaluate a possible effect of time since infection on risk of SEDI, we allowed the rate parameter for the fitted exponential distribution to change at 1 year post-infection and estimated the ratio of the post-1-year to pre-1-year rates (which we denote as the relative hazard). The model was estimated by maximum likelihood using the SAS NLMixed procedure (SAS Institute, Cary, NC, version 9.1). By definition, SEDI could not be detected at or before the time of the first genotyping, so this was considered to be a late entry time. The likelihood thus did not include any term for the period from infection to first genotype, reflecting certainty of no SEDI in that period. The exact time of SEDI was not known, so the likelihood reflected the chance of SEDI at any time between the last genotype not showing SEDI and the first one where it was detected. Because of the small number of events, we supplemented the standard Wald p-values with likelihood ratio p-values, based on comparing the primary model to one with no change in the rate parameter at 1 year post-infection.

Table 1: Incidence of Observed SEDI* in an Acute Infection Cohort by Year of Infection

	Year of Infection											Totals
	1	2	3	4	5	6	7	8	9	10	11	
SEDI* Event	6	1	2	0	0	0	0	0	0	0	0	9
Person-Years of Follow-up (overall - including time on treatment)	185.0	190.0	132.8	97.5	68.8	40.6	26.2	12.4	7.5	6.7	1.9	769.4
Incidence (overall) (per 100 person-years)	3.24	0.53	1.51	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.17
Person-years of Follow-up (only time off treatment)	153.8	151.7	108.4	86.3	58.9	38.4	25.2	11.4	7.0	6.7	1.9	649.7
Incidence (off treatment) (per 100 person-years)	3.90	0.66	1.85	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.44

*Sequentially expressed dual infections

Criteria for Selection: Participants have been on-study for at least 24 weeks and not on treatment at evaluation. Participants have baseline genotypes available or specimen at baseline (n = 100) available for genotyping. Specimens were selected for evaluation at four time points (5/28/2004, 4/30/2005, 3/5/2008, and 3/31/2009) if the participant had been off treatment for 24 weeks or remain treatment naive and of > 100.

Table 2: Cohort and SEDI Case Characteristics

	Cohort (n=266)	SEDI Cases (n=6)
Age	36.7	36.8
Male	96.6%	100%
Female	3.4%	0%
CD4 (average)		
Baseline	609	653
Assessment	609	528
Viral Load (average)		
Baseline	77,771	14,861
Assessment	73,743	84,370

Viral load in copies/ml

Results

266 recently infected persons were analyzed at 2 or more timepoints, representing 769.4 person-years of observation. The mean age was 36.7 and the cohort had been infected on average 103 days before their baseline genotype (Table 2). All viral sequences were subtype B. Highly divergent viruses appeared in 9 cases, representing an incidence density of 1.17/100 person-years. The observed incidence density of viral divergence was 3.24/100 person-years in the 1st year of infection, 0.53/100 person-years in the 2nd year, and 1.51/100 person-years in the 3rd year as seen in Table 1. No source partners have been identified for any of these cases.

Our primary model estimated an approximately 21-fold reduction in the risk of SEDI at times more than one year post-infection compared to times within one year of infection (relative hazard 0.05, 95% CI 0.01 to 0.39, p=0.0047). A likelihood ratio test of no difference between the two periods produced an even smaller p-value (p=0.0003). The estimated rate of SEDI was 4.4 per 100 person years (95% CI 2.2 to 9.0) in the first year following infection and was 0.2 per 100 person years beyond one year post-infection (95% CI 0.03 to 1.5).

The incidence densities observed in Table 1 differ from the estimates above because the estimates in Table 1 assume that the SEDI occurred when the second virus was first observed. The model-based analysis is likely to be more accurate since no a priori assumption was made about when the SEDI event occurred during the interval between first observation of the divergent virus and the last observed appearance of the first virus.

Discussion

The incidence of SEDI cases early in infection suggests that mechanisms partially blocking superinfection may develop over the first few years of infection. These mechanisms may include viral interference, depletion of target cells, or immune responses. Additional laboratory analysis is required to evaluate whether limited or localized superinfection has occurred in the absence of systemic overgrowth.

The dramatic estimated 21-fold reduction in risk of SEDI beyond one year post-infection could be due to biological factors as discussed above, but could also be caused by a phenomenon known as frailty selection. Decreasing risk over time could result from very heterogeneous levels of individual SEDI risk. The most susceptible persons would experience SEDI early, leaving only more resistant persons at later times, which would result in lower rates. We are in the process of gathering additional data, including detailed information about risk behavior over time, which may permit better assessment of the cause of decreasing SEDI risk with greater time since infection. In addition, ongoing routine sequencing and surveillance continues, as well as a sub-study to detect intermittent expression in the first 18 months. Sequencing of additional loci and increasing sensitivity for minor variant detection is also planned.

Acknowledgements

This work supported by NIH-NIAID; Program Project Grant P01 AI071713, U01 AI41531, and NIH R01 AI056988.