MAJOR ARTICLE

Quantifying Ongoing HIV-1 Exposure in HIV-1–Serodiscordant Couples to Identify Individuals With Potential Host Resistance to HIV-1

Romel D. Mackelprang,[1,2] Jared M. Baeten,[1,2,3] Deborah Donnell,[1,5] Connie Celum,[1,2,3] Carey Farquhar,[1,2,3] Guy de Bruyn,[7] Max Essex,[8,9] M. Juliana McElrath,[5,6] Edith Nakku-Joloba,[10] and Jairam R. Lingappa,[1,3,4] for the Partners in Prevention HSV/HIV Transmission Study Team

[1]Department of Global Health, [2]Department of Epidemiology, [3]Department of Medicine, and [4]Department of Pediatrics, University of Washington; [5]Statistical Center for HIV/AIDS Research and Prevention, and [6]The Vaccine and Infectious Disease Division, Fred Hutchinson Cancer Research Center, Seattle, Washington; [7]Perinatal HIV Research Unit, University of the Witwatersrand, South Africa; [8]Botswana-Harvard AIDS Institute Partnership, Gaborone, Botswana; [9]Harvard School of Public Health, Boston, Massachusetts; and [10]School of Public Health, Makerere University, Kampala, Uganda

Background. Immunogenetic correlates of resistance to HIV-1 in HIV-1–exposed seronegative (HESN) individuals with consistently high exposure may inform HIV-1 prevention strategies. We developed a novel approach for quantifying HIV-1 exposure to identify individuals remaining HIV-1 uninfected despite persistent high exposure.

Methods. We used longitudinal predictors of HIV-1 transmission in HIV-1 serodiscordant couples to score HIV-1 exposure and define HESN clusters with persistently high, low, and decreasing risk trajectories. The model was validated in an independent cohort of serodiscordant couples. We describe a statistical tool that can be applied to other HESN cohorts to identify individuals with high exposure to HIV-1.

Results. HIV-1 exposure was best quantified by frequency of unprotected sex with, plasma HIV-1 RNA levels among, and presence of genital ulcer disease among HIV-1–infected partners and by age, pregnancy status, herpes simplex virus 2 serostatus, and male circumcision status among HESN participants. Overall, 14% of HESN individuals persistently had high HIV-1 exposure and exhibited a declining incidence of HIV-1 infection over time.

Conclusions. A minority of HESN individuals from HIV-1–discordant couples had persistent high HIV-1 exposure over time. Decreasing incidence of infection in this group suggests these individuals were selected for resistance to HIV-1 and may be most appropriate for identifying biological correlates of natural host resistance to HIV-1 infection.

Investigators have sought to characterize correlates of resistance to human immunodeficiency virus type 1 (HIV-1) among individuals who have been exposed to HIV-1 yet remain seronegative, a population previously referred to by a diverse nomenclature but now commonly referred to as HIV-1–exposed seronegative (HESN) individuals [1, 2]. To date, however, only the CCR5-Δ32 mutation has consistently been associated with host resistance to HIV-1 [3], while observations regarding HIV-1–specific T cell responses, T-helper proliferation, interleukin 2 production, HIV-1–specific antibodies, immune activation, and CCL3L1 copy number variation have been inconsistent [4, 5].

One possible explanation for these disparate findings is that inaccurate quantification of HIV-1 exposure (ie, the likelihood that HIV-1 reaches target host cells) may result in misclassification of low-risk individuals as being highly exposed, leading to false associations with putative host resistance factors. For example, HESN individuals whose infected sex partners have initiated antiretroviral therapy (ART) [6–9]

Received 21 November 2011; accepted 9 April 2012.

Presented in part: 18th Conference on Retroviruses and Opportunistic Infections, #1029, Boston, Massachusetts, 1029, February 2011.

Correspondence: Romel D. Mackelprang, PhD, 325 Ninth Ave, Box 359927, Seattle, WA 98104 (romelm@u.washington.edu).

The Journal of Infectious Diseases 2012;206:1299–308

DOI: 10.1093/infdis/jis480

or who rarely have unprotected sex with their infected sex partners [10, 11] have lower risks of infection despite persistent sex with infected partners. In addition, absence of male circumcision [12–14], presence of other sexually transmitted infections [15, 16], and current pregnancy [17] have repeatedly been associated with HIV-1 transmission. These associations may be attributable to changes in the infectiousness of source partners or may reflect their influence on the likelihood that HIV-1 reaches target host cells in uninfected partners. Indeed, HIV-1 exposure characteristics may collectively modify infection risk by up to 300-fold [18]. Furthermore, because these characteristics are dynamic, cross-sectional measurement may inadequately capture ongoing levels of exposure. Thus, quantifying longitudinal HIV-1 exposure in biological studies of resistance could reduce exposure misclassification and, ultimately, improve precision.

To evaluate how factors quantifying HIV-1 exposure can predict HIV-1 acquisition risk, we modeled HIV-1 exposure using longitudinal data from a cohort of HIV-serodiscordant couples and validated the model in an independent cohort. We used this model to identify a subset of HESN partners with persistently high HIV-1 exposure who exhibited a declining HIV-1 infection incidence, suggesting selection for individuals with resistance to HIV-1. Thus, we propose that this approach to quantifying longitudinal HIV-1 exposure can identify individuals most likely to be enriched for biological factors mediating host resistance to HIV-1.

METHODS

Cohorts and Study Procedures

HIV-1 exposure scores were developed using data from a randomized clinical trial of 3408 HIV-serodiscordant heterosexual African couples that evaluated the efficacy of acyclovir in preventing HIV-1 transmission over 12–24 months of quarterly follow-up, as previously described [19]. All HIV-1–infected partners were herpes simplex virus 2 (HSV-2) seropositive, had CD4 cell counts $\geq$250 cells/mm^3, and were not taking ART at enrollment. A separate validation cohort of 485 HIV-serodiscordant couples from Kampala, Uganda, and Soweto, South Africa, was followed quarterly for 1 year in an observational study of HIV-1 transmission. This secondary cohort did not have the same HSV-2 or CD4 cell count eligibility criteria as the primary cohort. Both studies determined HIV-1 infection by serologic analysis, and infection dates were estimated using reverse-transcription polymerase chain reaction (RT-PCR) analysis of the HIV-1 RNA level in preseroconversion plasma [20]. Genetic sequencing of *env* and *gag* gene regions of HIV-1 obtained from plasma specimens from both partners was used to assess transmission linkage within the partnership [19, 21]. Both studies were approved by human subjects research committees at the University of Washington and all local study sites and affiliated institutions. All participants provided written informed consent.

Definition and Correlates of HIV-1 Exposure

HIV-1 exposure was conceptualized as HIV-1 reaching host target cells following heterosexual contact with an HIV-infected partner and was defined by correlates of HIV-1 infectivity and susceptibility. Specifically, we evaluated plasma HIV-1 RNA levels, pregnancy, and genital ulcer disease among infected partners [8,9,22], unprotected sex between partners [10,11], and sex, age, male circumcision [12–14], HSV-2 serostatus, pregnancy [17], and genital ulcer disease [15,16] among HESN partners. We did not include CD4 cell counts or ART use by the HIV-1–infected partner since effects of these factors are correlated with HIV-1 RNA levels. Age, sex, and male circumcision data were obtained at enrollment. Plasma HIV-1 RNA levels were determined using the COBAS AmpliPrep/COBAS TaqMan HIV-1 RNA assay, version 1.0 (Roche Diagnostics, Indianapolis, IN), with a limit of quantification of 240 copies/mL, on plasma collected from infected partners at the enrollment visit; the 3-, 6-, 9-, and 12-month visits; and study exit.

We created a time-dependent dichotomous variable based on unprotected sex reported by either partner since the last quarterly visit. The genital ulcer disease variable included self-report of genital ulcer disease or ulcer diagnosis upon quarterly examination. Pregnancy was assessed using urine tests for HIV-1–infected women and using self-report and optional urine tests for HESN women. Pregnancy duration was defined as the time from the last menstrual period preceding pregnancy to delivery or pregnancy loss.

HIV-1 Outcomes

Infection dates for HIV-1 seroconverters without a positive HIV-1 RNA RT-PCR result before seroconversion were estimated as the midpoint between the last seronegative visit and the first seropositive visit, or 45 days before the first seropositive visit for individuals who missed a quarterly follow-up visit and thus had a large gap between visits. For seroconverters with a positive RT-PCR result prior to seroconversion, infection dates were estimated as 17 days before the visit with first RT-PCR–positive result [23, 24].

HIV-1 Exposure Quantification

To quantify HIV-1 exposure among HESN participants, we determined a best-fitting Cox proportional hazards model with time-dependent covariates for predicting HIV-1 infection, using backward variable selection based on Akaike's information criterion, and we used regression coefficients and individual covariates to estimate visit-specific exposure scores. The Cox model outcome was time from enrollment to acquisition of HIV-1 infections that were genetically linked to the source partner's virus. Seroconverters who acquired HIV-1 from outside partnerships

were censored at their estimated date of infection, and those who remained uninfected were censored at their last HIV-1 test. Potential predictors included the last known value before each HIV-1 test for each covariate described above. Validation of the model using the secondary cohort was performed using receiver operating characteristic (ROC) curves [25].

HIV-1 exposure scores (η_{it}) at each visit t were determined for every HESN participant i by use of linear predictors from the final model. Because the Cox model is a relative risk model, linear predictors were normalized to the sample average so the average exposure score was 0, with scores >0 representing greater than average risk. Specifically, the exposure score for participant i at study visit t was calculated as follows:

$$\text{exposure score}_{it} = \eta_{it} = \beta X_{it} - \left(\frac{1}{n}\right)\sum_{j}\beta X_j$$

where X and β represent covariates and model coefficients, respectively, and X_j represents the covariate value for the *jth* of n participants. The vector of exposure scores for an individual i over follow-up $(\eta_i = [\eta_{i1}, \eta_{i2}, \ldots, \eta_{ik}])$ is referred to as their "risk trajectory."

Finally, we subdivided participants into clusters with similar HIV-1 exposure trajectories, using K-means cluster analysis that can handle missing values [26]. Depending on the number of prespecified clusters, this approach typically identifies a cluster with persistent high exposure, 1 or more clusters with persistently lower levels of HIV-1 exposure, and a cluster whose exposure scores decreased dramatically over follow-up. We plotted hazard functions for these risk groups to evaluate empirical changes in infection risk [27].

Simulating Effects of HIV-1 Exposure in Studies of Resistance

We used simulations to demonstrate effects of selecting low exposure HESN controls when studying a potential correlate of resistance in scenarios in which we assumed different associations between exposure and a hypothetical host factor and different associations between the hypothetical host factor and HIV-1 acquisition. For each scenario, we simulated 100 populations of 3400 participants with similar exposure score distributions, HIV-1 incidences, and associations between exposure scores and infection as in the primary cohort. Next, we simulated the distribution of a continuous host factor in each population on the basis of assumed true relationships of the host factor with both exposure scores and infection. Finally, we evaluated observed associations between the hypothetical host factor and HIV-1 infection when HESN controls were selected either randomly or by exposure score matching to cases.

Supplemental R Program for Determining Exposure Scores

We developed a free R package for determining HIV-1 exposure scores that can be applied to other cohorts (Supplementary Materials).

RESULTS

Baseline Characteristics

Of 3408 HESN participants, 3321 had ≥1 follow-up HIV-1 test. A total of 2236 (67.3%) tested HESN participants were male, and, of these, 1336 (59.7%) were circumcised. Median ages among uninfected females and males were 31 years (interquartile range [IQR], 26–38 years) and 35 years (IQR, 30–42 years), respectively. At baseline, ≥1 partner in 1175 couples (35.4%) reported unprotected sex in the previous month, and 2261 HESN participants (68%) were HSV-2 seropositive. Furthermore, 780 HIV-1–infected partners (23%) and 344 HIV-1–uninfected partners (10%) had symptomatic genital ulcer disease. Among HIV-1–infected partners, the median baseline plasma HIV-1 RNA level was 4.1 $\log_{10}$ copies/mL (IQR, 3.3–4.7 $\log_{10}$ copies/mL).

HIV-1 Infection Incidence

Seroconversion was detected in 151 participants who were initially HESN (4.5%), of whom 24 were HIV-1 RNA RT-PCR positive at enrollment and were excluded from the analysis. Of the remaining 127 seroconverters, HIV-1 in 86 (67.7%) was genetically linked to HIV-1 in their infected partner, as revealed by viral sequencing [21]. The overall incidence of linked infections was 1.7 cases/100 person-years and was greater in the first year of follow-up than in the second (2.0 vs 1.2 cases/100 person-years; $P = .04$).

HIV-1 Exposure Scores

By use of backward variable selection, the best-fitting model for HIV-1 acquisition included unprotected sex with, HIV-1 RNA load of, and symptomatic genital ulcer disease for the infected partner and HSV-2 serostatus, current pregnancy, sex, age, and male circumcision of the uninfected partner (Table 1). Evaluation of Schoenfeld residuals did not suggest any time-varying effects. We used this model to calculate visit-specific exposure scores for HESN participants on the basis of the product of regression coefficients and the participant's covariates, and we normalized these to the average exposure score across the full cohort. Exposure scores ranged from −3.6 to 4.7, with a score of 0 representing the average exposure and a 1-unit increase indicating an exp(1) = 2.7-fold increased risk of infection.

Longitudinal Changes in HIV-1 Exposure

Because of variation in time-dependent predictors, HIV-1 exposure scores varied across study visits. For instance, among 1715 (52%) participants who reported unprotected sex at ≥1 visit, 938 (47%) reported unprotected sex at ≤25% of their visits. Among 2764 HESN participants whose HIV-1–infected partner had detectable HIV-1 plasma RNA at baseline, 241 (8.5%) had undetectable plasma HIV-1 levels by the end of follow-up, and 163 HIV-1–infected partners who always had a

Table 1. Multivariable Hazard Ratios and Regression Coefficients From the Best-Fitting Cox Proportional Hazards Model Used to Estimate Human Immunodeficiency Type 1 (HIV-1) Exposure Scores

Characteristic	Hazard Ratio	95% CI	*P*	Regression Coefficient
Couple				
Any unprotected sex[a,b]	4.2	(2.7–6.6)	<.001	1.4
Seropositive partner				
HIV-1 RNA load (per 1 $\log_{10}$ increase)[a]	2.7	(2.1–3.5)	<.001	1.0
Genital ulcer disease[a]	1.6	(.9–2.8)	.08	0.5
Seronegative partner				
Pregnancy[a]	1.8	(.9–3.8)	.12	0.6
HSV-2 seropositive	2.1	(1.2–3.8)	.01	0.8
Male circumcision	0.6	(.3–1.1)	.10	−0.5
Age (10 y increase)[c]				
Among females	0.4	(.2–0.8)		−0.8
Among males	0.8	(.5–1.1)		−0.3
Female vs male sex[c]				
At 25 y of age	0.9	(.5–1.8)		−0.1
At 40 y of age	0.5	(.2–1.01)		−0.8

The best-fitting model was determined using backward variable selection with Akaike's information criterion as the stopping rule, which resulted in retention of some covariates with $P > .05$.

Abbreviations: CI, confidence interval; HSV-2, herpes simplex virus 2.

[a] Frequency of unprotected sex, plasma HIV-1 RNA levels, genital ulcer disease, and pregnancy were modeled as time-dependent variables. Baseline measurements from enrollment were used for all other covariates.

[b] Unprotected sex was indicated if either partner reported at least 1 occurrence of sexual intercourse without using a condom since the previous quarterly visit.

[c] Evidence of statistical interaction between the sex and age of the HIV-1–seronegative partner ($P = .1$) suggests that increasing age results in more substantial decrease in HIV-1 acquisition risk among females than males. The coefficients associated with age, sex, and the interaction were −0.2, 1.0, and −0.4, respectively.

detectable viral load experienced a decrease of ≥1 $\log_{10}$ copies. Of these 404 HIV-1–infected participants, 124 (31%) began using ART during the study. During follow-up, 1230 HIV-1–infected partners (37%) had genital ulcer disease by self-report or physical examination at ≥1 visit; however, ulcers were only found at 1946 (26%) of 7500 study visits attended by these participants. Finally, of 1873 visits attended by 293 uninfected women who were pregnant at any time during the study, pregnancy was documented at 699 (37%). Overall, only 51% of participants in the highest exposure score quintile at baseline remained in the highest exposure quintile at the 3-month follow-up visit.

HESN Clusters With Persistent Levels of HIV-1 Exposure

Despite variability across individual visits, participants could be divided into clusters with persistent high exposure scores (475 [14%]), stable lower risk scores (2595 [79%]), or decreasing exposure scores (214 [7%]) (Supplementary Figure 1). Compared with participants with low exposure scores, participants in the highest exposure group exhibited riskier characteristics (Table 2). Specifically, high-exposure participants were more likely to report at any time during the study that they had unprotected sex with their study partner (75% vs 46%; $P < .001$), and their HIV-1–infected partners had higher mean plasma HIV-1 RNA levels (5.0 vs 3.9 $\log_{10}$ copies/mL; $P < .001$). Furthermore, these participants were younger (mean age, 29.5 vs 34.9 years; $P < .001$) and were more likely to be female (50% vs 29%; $P < .001$). Among male HESN participants, the highest-exposure group was less likely to be circumcised (31% vs 59%; $P < .001$).

Participants in the highest-exposure group had a 6.9-fold increased risk of infection than the lower exposure group on the basis of median exposure scores (1.7 [IQR, 1.5–2.1] vs −0.2 [IQR, −1.0 to 0.5]) and had similar median exposure scores as participants who acquired HIV-1 (Figure 1). Furthermore, this group had the highest incidence of HIV-1 acquisition, with 49 individuals (10%) acquiring HIV-1 during follow-up, compared with only 28 (1%) in the lower-exposure group. Empirical plots showing smoothed hazards of infection among exposure clusters suggested that the risk of infection decreased over time among participants in the highest-risk group but remained constant among participants with lower exposure scores (Figure 2).

Participants in the cluster with substantial decreases in exposure over follow-up started at baseline with a median exposure risk score of 1.4 (IQR, 0.9–3.0) but had much lower scores across all subsequent visits (median, 0.4 [IQR, −0.5 to 0.9]). This drop in exposure was principally due to cessation of unprotected sex or to the HIV-1–infected partner's HIV-1 RNA levels decreasing after initiation of ART, with 39% of infected partners of participants in the decreasing exposure group reporting ART use during the study, compared with 13% and 7% in the highest- and lower-risk groups, respectively ($P < .001$).

Evaluation of Simplified and HESN-Only HIV-1 Exposure Score Models

To compare the best fitting longitudinal model to models that can be applied in cohorts with less data, we also evaluated a simplified model that included only baseline HIV-1 RNA levels of infected partners and longitudinal unprotected sex, along with age, sex, and male circumcision status of the HESN partner, and a HESN-only model that included predictors from the HESN partner (unprotected sex, sex, age, and male circumcision status) without behavioral or clinical data from the HIV-1–infected partner. Compared with the best-fitting model, mean individual exposure scores from the simplified model discriminated seroconverters from nonseroconverters with a high degree of sensitivity and specificity, as measured

Table 2. Characteristics of Longitudinal Exposure Risk Groups

Characteristic	HIV-1 Exposure Score Group				
	Highest (n = 475)	Decreasing (n = 214)	P^a	Lower (n = 2595)	P^a
Mean HIV-1 exposure score during follow-up[b]	1.7 (1.5–2.0)	0.4 (–0.5 to 0.9)	<.001	–0.2 (–1.0 to 0.5)	<.001
Linked HIV-1 transmission[b]	49 (10)	7 (3)	<.05	28 (1)	<.05
Couples					
Married to study partner	386 (81)	154 (72)	<.05	1951 (75)	<.05
Unprotected sex during study[b]	357 (75)	139 (65)	<.05	1203 (46)	<.001
Relationship duration, y	4.1 (1.8–7.5)	5.0 (2.4–9.7)	<.05	4.8 (2.0–10.1)	<.05
HIV-1–seronegative partner					
Age, y	30 (25–35)	32 (27–40)	<.001	35 (29–42)	<.001
Female sex	238 (50)	71 (33)	<.001	761 (29)	<.001
Male circumcision	74 (31)	68 (48)	<.05	1077 (59)	<.001
HSV-2 seropositive	417 (88)	146 (68)	<.05	1680 (65)	<.001
Pregnant during study[b]	111 (47)	17 (24)	<.05	163 (21)	<.001
HIV-1–seropositive partner					
Mean plasma HIV-1 RNA level,[b,c] $\log_{10}$ copies/mL	5.0 (4.6–5.3)	4.0 (3.6–4.5)	<.001	3.9 (3.2–4.4)	<.001
Mean CD4 cell count, copies/mm^3[b,c]	371 (291–488)	373 (298–531)		466 (351–626)	<.001
Initiated ART during study[b,d]	64 (13)	83 (39)	<.001	181 (7)	<.001
Time of ART initiation after enrollment, mo[b,d,e]	15 (15–21)	9 (9–12)	<.001	15 (9–18)	<.05
Genital ulcer disease during study[b]	224 (47)	95 (44)		905 (35)	<.001

Data are no. (%), for categorical variables, and medians (interquartile ranges), for continuous variables. Numbers may not sum to the total number of participants included in the study, because of missing data.

Abbreviations: ART, antiretroviral therapy; HIV-1, human immunodeficiency virus type1.

[a] Values denote results of comparisons of groups with decreasing or lower HIV-1 exposure scores to those with the highest exposure scores. χ^2 tests were used for categorical variables, and *t* tests were used for continuous variables.

[b] HIV-1 exposure scores, plasma HIV-1 RNA levels, and CD4 cell counts are provided as the mean value across all study visits. HIV-1 transmission, unprotected sex, pregnancy, and genital ulcer disease variables indicate if the condition was observed at any visit during follow-up. Baseline measurements from enrollment were used for all other covariates.

[c] Mean values of plasma HIV-1 RNA levels and CD4 cell counts for each individual across all study visits.

[d] Referrals for ART initiation were based on national guidelines at the time of the study.

[e] HIV-1–infected partners were eligible for the study if they had CD4 cell counts >250 copies/mm^3 and were not receiving ART at enrollment following national guidelines.

by areas under the ROC curve (AUCs) of 0.87 versus 0.85. Furthermore, 71% of participants identified in the highest-exposure cluster identified using the best-fitting model were in the highest-risk cluster identified by the simplified model. Only including baseline unprotected sex further reduced the AUC to 0.82. The discriminatory power of a model that included predictors from only the HESN participant was not as strong as either the best-fitting model or the simplified model, as indicated by an AUC of 0.71, with only 36% of highest-risk participants identified through the best-fitting model being captured as high risk in the model that was based on HESN participant data only.

Model Validation

To validate use of exposure scores for discriminating HIV-1 seroconverters from HESN participants, we used the simplified model described above to determine exposure scores in a second cohort of 485 HIV-1 serodiscordant couples. Although the second cohort was recruited similarly to the primary cohort, the second study only evaluated plasma HIV-1 RNA levels at baseline, hence necessitating validation with the simplified model only. Mean individual exposure scores generated by the simplified model applied to this second cohort discriminated seroconverters from nonseroconverters with a high degree of sensitivity and specificity, as measured by an AUC of 0.81, which was similar to the AUC of 0.85 achieved by the simplified model in the primary cohort (Figure 3). The simplified model could also be used in the second cohort to identify a highest risk cluster composed of 48 participants (10.4%), which also showed a decreasing hazard of HIV-1 infection over time.

Simulating Effects of HIV-1 Exposure in Studies of Resistance

We used data from our primary cohort to evaluate potential effects of selecting HESN controls with low exposure when studying a hypothetical biological factor that does or does not correlate with host resistance to HIV-1. Scenario 1 assumes that increased exposure is associated with increased levels of a

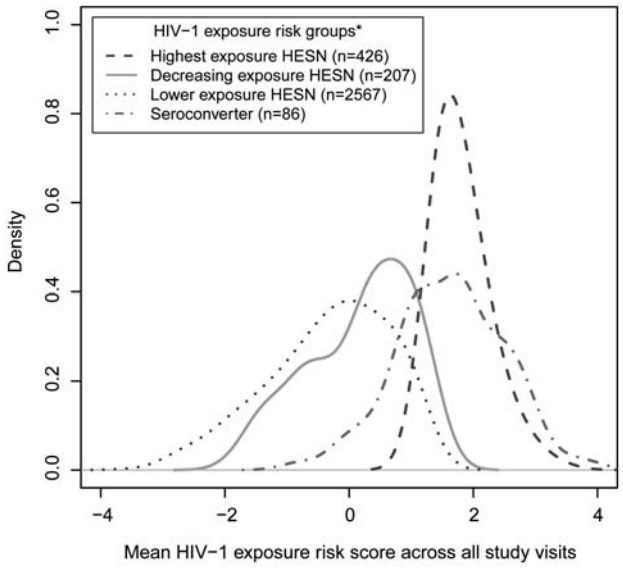

Figure 1. Smoothed density curves representing human immunodeficiency virus type 1 (HIV-1) exposure score distributions for all HIV-1 seroconverters and HIV-1–exposed seronegative participants from the highest, lower, and decreasing exposure risk groups. Longitudinal HIV-1 exposure scores were quantified using time-dependent predictors (unprotected sex, plasma HIV-1 RNA levels, and symptomatic genital ulcer disease in HIV-1–infected partners and age, pregnancy, herpes simplex virus 2 serostatus, and male circumcision in HIV-1–exposed seronegative participants), with a 1-unit increase representing a exp(1)=2.7-fold increased risk of HIV-1 acquisition. HIV-1 exposure risk groups were based on individual exposure score trajectories over time and were created using longitudinal K-means cluster analysis. Area under the kernel density curves between 2 HIV-1 exposure risk scores represents the probability that exposure risk scores for individuals in the respective participant subgroup fell between those 2 values of the exposure score.

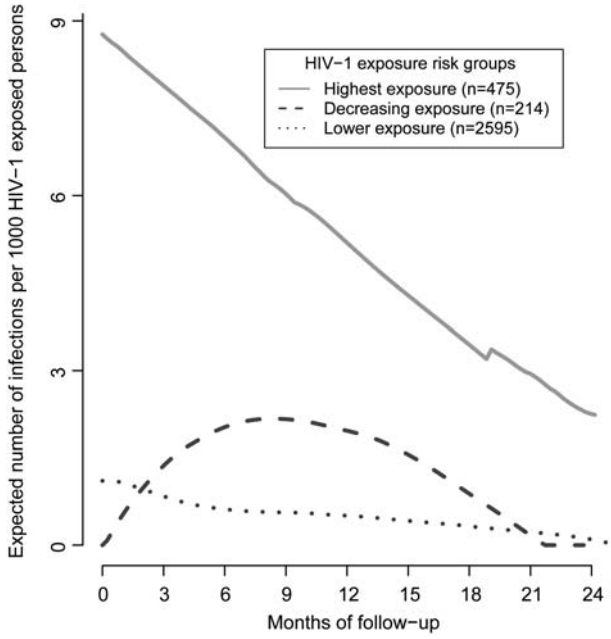

Figure 2. Empirical hazard functions for human immunodeficiency virus type 1 (HIV-1) acquisition among HIV-1 exposure score risk groups, determined by clustering initially HIV-1-exposed seronegative individuals into homogenous groups on the basis of their longitudinal HIV-1 exposure trajectories. Hazard rates represent the instantaneous risk of HIV-1 acquisition at time *t* conditional on survival until time *t* or later.

hypothetical host factor and that this host factor is not associated with infection (Figure 4*A*). If controls are selected at random (eg, not on the basis of HIV-1 exposure), a spurious relationship between the host factor and HIV-1 acquisition is observed (mean difference, 1.7; $P < .001$) that reflects confounding of HIV-1 exposure on HIV-1 acquisition (Figure 4*B*). The correct relationship between the hypothetical host factor and HIV-1 acquisition is observed once the level of HIV-1 exposure is controlled by matching HESN participants to HIV-1 seroconverters by exposure score (observed mean difference, 0; $P = .6$) (Figure 4*C*). Scenario 2 assumed that increased exposure is associated with increased host factor levels but that the host factor was associated with infection, with the mean level being 2 units lower among HIV-1 seroconverters (Figure 4*D*). Here, randomly selecting HESN controls results in a false-negative association (observed mean difference, 0.2; $P = .4$) (Figure 4*E*). Once again, this false observation is rectified by matching cases and controls by exposure levels, revealing the true association between the hypothetical host factor and HIV-1 acquisition (mean difference, −1.97; $P < .001$) (Figure 4*F*).

DISCUSSION

Our HIV-1 exposure scores for HESN partners from HIV-1 serodiscordant couples can prioritize participants for studying biological correlates of resistance to HIV-1 and can be used to adjust regression models when conducting future studies or revisiting past analyses. By using this approach, we identified a small subset of HESN participants (14%) with persistently elevated HIV-1 exposure during follow-up who, consequentially, had the highest HIV-1 acquisition rates. However, although the instantaneous risk of acquiring HIV-1 in this subgroup of HESN partners was highest early during follow-up, the incidence of HIV-1 infection declined markedly over follow-up. This decreasing HIV-1 infection risk is consistent with observations that the HIV-1 infection incidence among Kenyan commercial sex workers decreased with increased duration of sex work [28]. A plausible interpretation is that declining HIV-1 incidence despite persistent high HIV-1 exposure in our cohort reflects selection for HIV-1–resistant individuals. Our approach may therefore permit identification of HESN individuals who are most likely to yield biological correlates

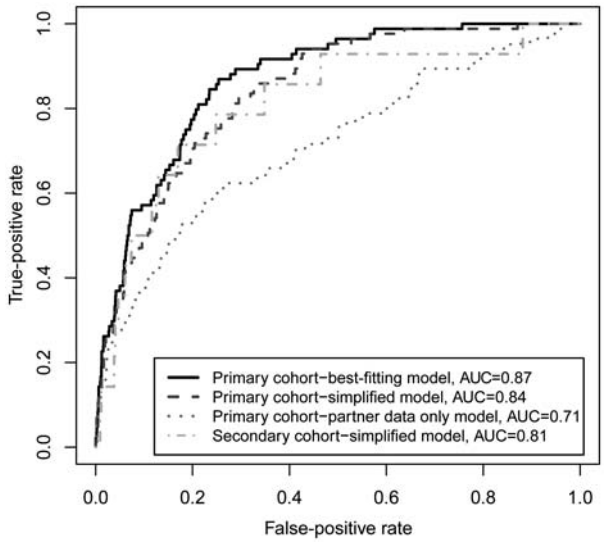

Figure 3. Receiver operating characteristic (ROC) curves comparing the ability of an individual's average human immunodeficiency virus type 1 (HIV-1) exposure score across all study visits to discriminate HIV-1 acquisition risk for participants in the primary and secondary cohorts. Models for HIV-1 acquisition were developed with primary cohort of 3408 initially seronegative partners from HIV-1–discordant couples in the Partners in Prevention HSV/HIV Transmission Study. The final best-fitting Cox proportional hazards model for HIV-1 acquisition included unprotected sex with, HIV-1 RNA level of, and genital ulcer disease for the infected partner and herpes simplex virus 2 (HSV-2) serostatus, pregnancy, sex, age, and male circumcision for the uninfected partner. The reduced model included the same variables as the primary model but used baseline rather than longitudinal plasma HIV-1 RNA levels. The secondary cohort included 485 seronegative partners from HIV-1 discordant couples in the Couples Observational Study. Only baseline HIV-1 RNA levels for infected partners were available for this cohort, so ROC curves were generated for a reduced model that included baseline but not time-varying HIV-1 RNA concentrations. Abbreviation: AUC, area under the ROC curve.

when studying host resistance to HIV-1 and may limit false-positive and spurious results.

Our model used established predictors of HIV-1 acquisition, with plasma HIV-1 RNA levels of infected partners and unprotected sex having the strongest effects. Exposure scores were dynamic, and our longitudinal model had greater sensitivity and specificity than a cross-sectional model. Yet, predictive power for HIV-1 infection remained high when the baseline plasma HIV-1 RNA load was substituted for longitudinal HIV-1 levels while keeping longitudinal information for other variables. This likely reflects high concordance in HIV-1 RNA levels over time in the absence of ART initiation.

A strength of this analysis was use of epidemiologic and clinical data for both sex partners, which substantially improved exposure quantification. Confirmation of HIV-1 transmission linkage by use of viral genetic sequences further improved model precision (Supplementary Figure 2). We also demonstrated the model's predictive capacity in an independently recruited validation cohort of couples with stable HIV-1 serodiscordance. The validity of our approach may be limited in epidemiologic contexts in which little is known about HIV-1–infected partners, such as cohorts of commercial sex workers or high-risk men who have sex with men. This is supported by reduced predictive capacity when only using data from HESN partners. Nevertheless, each cohort includes unique data that may improve the performance of exposure models. To facilitate such evaluations, we have created a freely available program for quantifying HIV-1 exposure in diverse epidemiologic contexts.

Unsystematic quantification of HIV-1 exposure has likely contributed to widely disparate correlates of resistance. This is demonstrated in our simulations showing that inadequate control of exposure may cause false-positive or false-negative findings in host factor studies if exposure is associated with that factor. This situation could arise through 2 mechanisms. First, HIV-1 exposure may be directly associated with an immunologic factor via a direct biological relationship. For example, HIV-specific interferon γ cytotoxic T lymphocyte responses among Kenyan HESN commercial sex workers were lost upon cessation of sex work [29]. Furthermore, changes in memory and activated T cells from HESN partners in serodiscordant couples are strongly correlated with plasma viral loads of the infected partner [30]. HIV-1 exposure levels may also be indirectly associated with a correlate of resistance. For instance, a protective phenotype such as CCR5-Δ32 may be enriched in highly exposed HESN individuals since highly exposed persons without that mutation are likely to become infected early [31]. Last, while the HIV-1 exposure scores will be useful for selecting HESN individuals with the highest levels of exposure for studies of immune correlates of protection, it will be important to adjust for additional potential confounding variables (eg, persistent inflammation from genital herpes [32]) on the basis of populations and cofactors of each study.

Our analysis has several advantages over a recently reported mathematical model of HIV-1 exposure–based risk [33]. First, our model uses actual data from our primary cohort, whereas the multipliers in the Bernoulli model came from the published literature. The accuracy of that model assumes that multiplier effects are independent and are not distorted by inclusion of other risk factors in the model, which may be incorrect for highly correlated factors, such as an infected partner's HIV-1 load and disease stage. Exposure scores derived from our model are likely more accurate because we accounted for the interdependence of predictive factors by adjusting the effects of each predictor for other variables in the model. Second, the usefulness of our model is further supported by its capacity to discriminate HIV-1 acquisition risk through validation in an independently recruited cohort.

In summary, our approach to quantifying longitudinal HIV-1 exposure risk may improve the sensitivity and specificity of studies seeking to identify biological correlates of HIV-1

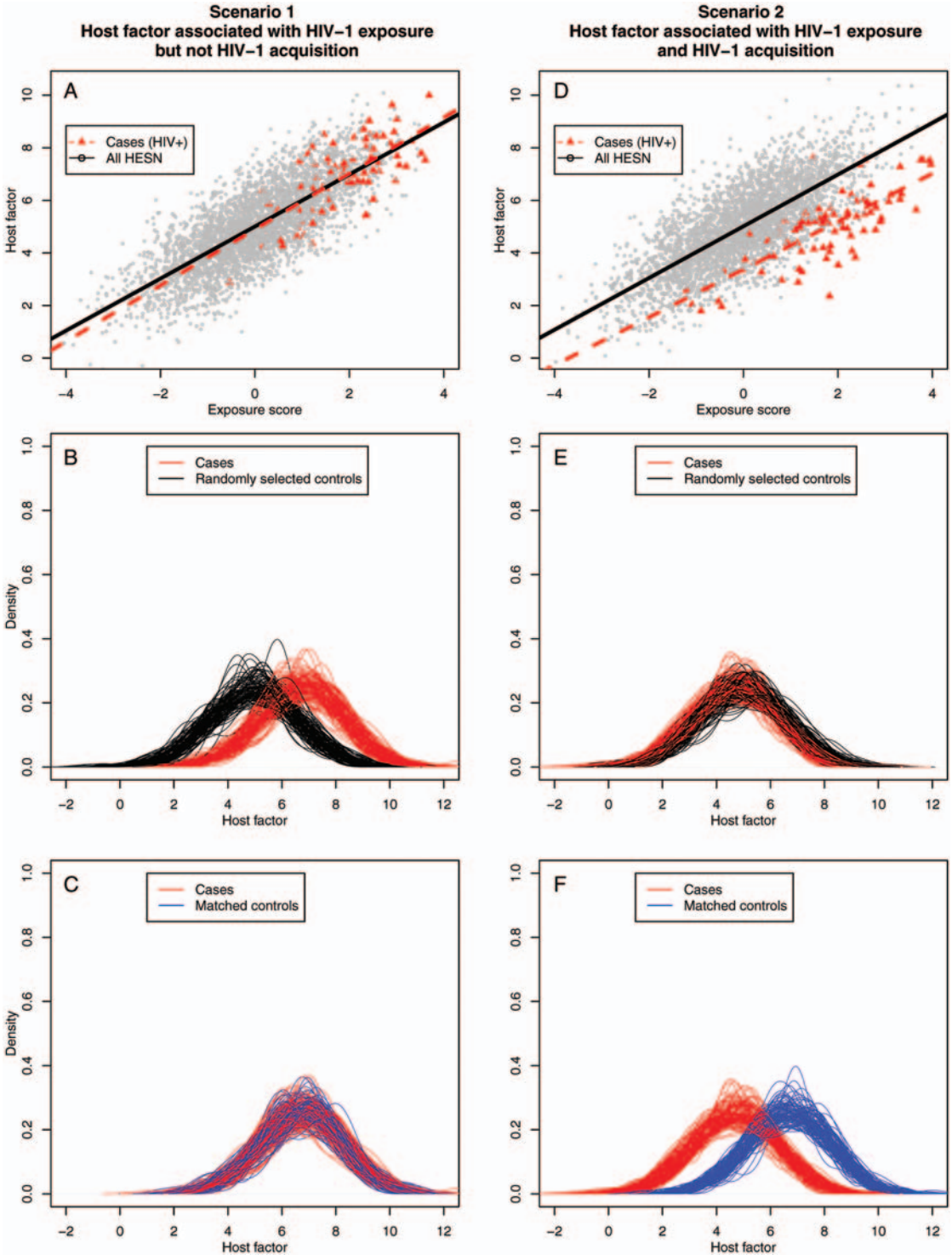

Figure 4. Simulations demonstrating potential biases when evaluating potential correlates of human immunodeficiency virus type 1 (HIV-1) resistance if HIV-1 exposure is not considered. Scenario 1 assumes that a 1-unit increase in HIV-1 exposure score is associated with a 1-unit increase in a continuous hypothetical host factor and that the host factor is not associated with HIV-1 acquisition. *A*, True relationship of exposure score and HIV-1 exposure score with no difference in host factor for seroconverters and HIV-1–exposed seronegative (HESN) individuals. *B*, False-positive association of HESN with elevation in the hypothetical host factor due to random selection of controls without regard to HIV-1 exposure levels. *C*, True-negative association revealed by selecting controls with similar exposure scores as seroconverters. Scenario 2 assumes that a 1-unit increase in HIV-1 exposure score is associated with a 1-unit increase in a continuous hypothetical host factor and that the average level of the continuous host factor was 2 units lower among seroconverters than among HESN individuals of the same exposure level. *D*, True relationship of exposure score and HIV-1 exposure score with a 2-unit difference in host factor for seroconverters and HESN individuals. *E*, False-negative association of HESN with the hypothetical host factor due to random selection of controls. *F*, True-positive association revealed by selecting controls with similar exposure scores as seroconverters.

resistance [1]. Our approach to estimating HIV-1 exposure using longitudinal data from both partners in HIV-1–serodiscordant couples provides an objective tool to identify subsets of HESN individuals to target for identification of host factors protecting against HIV-1.

Supplementary Data

Supplementary materials are available at *The Journal of Infectious Diseases* online (http://jid.oxfordjournals.org/). Supplementary materials consist of data provided by the author that are published to benefit the reader. The posted materials are not copyedited. The contents of all supplementary data are the sole responsibility of the authors. Questions or messages regarding errors should be addressed to the author.

Notes

Acknowledgments. The Partners in Prevention HSV/HIV Transmission Study Team: University of Washington Coordinating Center and Central Laboratories, Seattle: Connie Celum (principal investigator), Anna Wald (protocol cochair), Jairam Lingappa (medical director), Jared M. Baeten, Mary Campbell, Lawrence Corey, Robert W. Coombs, James P. Hughes, Amalia Magaret, M. Juliana McElrath, Rhoda Morrow, and James I. Mullins. Study sites and site principal investigators: in Cape Town, South Africa, David Coetzee (University of Cape Town); in Eldoret, Kenya, Kenneth Fife and Edwin Were (Moi University and Indiana University); in Gaborone, Botswana, Max Essex and Joseph Makhema (Botswana Harvard Partnership); in Kampala, Uganda, Elly Katabira and Allan Ronald (Infectious Disease Institute, Makerere University); in Kigali, Rwanda, Susan Allen, Kayitesi Kayitenkore, and Etienne Karita (Rwanda Zambia HIV Research Group and Emory University); in Kisumu, Kenya, Elizabeth Bukusi and Craig Cohen (Kenya Medical Research Institute and University of California–San Francisco); in Kitwe, Zambia, Susan Allen and William Kanweka (Rwanda Zambia HIV Research Group and Emory University); in Lusaka, Zambia, Susan Allen and Bellington Vwalika (Rwanda Zambia HIV Research Group and Emory University); in Moshi, Tanzania, Saidi Kapiga and Rachel Manongi (Kilimanjaro Christian Medical College and Harvard University); in Nairobi, Kenya, Carey Farquhar, Grace John-Stewart, and James Kiarie (University of Nairobi and University of Washington); in Ndola, Zambia: Susan Allen and Mubiana Inambao (Rwanda Zambia HIV Research Group and Emory University); in Orange Farm, South Africa, Sinead Delany-Moretlwe and Helen Rees (Reproductive Health Research Unit and University of the Witwatersrand); in Soweto, South Africa, Guy de Bruyn, Glenda Gray, and James McIntyre (Perinatal HIV Research Unit, University of the Witwatersrand); in Thika, Kenya, Nelly Rwamba Mugo (University of Nairobi and University of Washington). Data management was provided by DF/Net Research (Seattle), and site laboratory oversight was provided by Contract Lab Services (University of the Witwatersrand, Johannesburg, South Africa).

Disclaimer. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Financial support. This work was supported by the Bill and Melinda Gates Foundation (grants 26469 and 41185) and the NIH/NIAID (grant AI073115).

Potential conflicts of interest. All authors: No reported conflicts.

All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed.

References

1. Meyers AFA, Fowke KR. International symposium on natural immunity to HIV: a gathering of the HIV-exposed seronegative clan. J Infect Dis **2010**; 202(Suppl):S327–8.
2. Young JM, Turpin JA, Musib R, Sharma OK. Outcomes of a national institute of allergy and infectious diseases workshop on understanding HIV-exposed but seronegative individuals. AIDS Res Hum Retroviruses **2011**; 27:737–43.
3. Huang Y, Paxton WA, Wolinsky SM, et al. The role of a mutant CCR5 allele in HIV-1 transmission and disease progression. Nat Med **1996**; 2:1240–3.
4. Lederman MM, Alter G, Daskalakis DC, et al. Determinants of protection among HIV-exposed seronegative persons: an overview. J Infect Dis **2010**; 202(Suppl):S333–8.
5. Restrepo C, Rallón NI, Carrillo J, et al. Host factors involved in low susceptibility to HIV infection. AIDS reviews **2011**; 13:30–40.
6. Cohen MS, Chen YQ, McCauley M, et al. Prevention of HIV-1 infection with early antiretroviral therapy. N Engl J Med **2011**; 365:493–505.
7. Donnell D, Baeten JM, Kiarie J, et al. Heterosexual HIV-1 transmission after initiation of antiretroviral therapy: a prospective cohort analysis. Lancet **2010**; 375:2092–8.
8. Lingappa JR, Hughes JP, Wang RS, et al. Estimating the impact of plasma HIV-1 RNA reductions on heterosexual HIV-1 transmission risk. PLoS one **2010**; 5:e12598.
9. 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.
10. Wawer MJ, Gray RH, Sewankambo NK, et al. Rates of HIV-1 transmission per coital act, by stage of HIV-1 infection, in Rakai, Uganda. J Infect Dis **2005**; 191:1403–9.
11. Hughes JP, Baeten JM, Lingappa JR, et al. Determinants of per-coital-act HIV-1 infectivity among African HIV-1-serodiscordant couples. J Infect Dis **2012**; 205:358–65.
12. Auvert B, Taljaard D, Lagarde E, et al. Randomized, controlled intervention trial of male circumcision for reduction of HIV infection risk: the ANRS 1265 Trial. PLoS Med **2005**; 2:e298.
13. Bailey RC, Moses S, Parker CB, et al. Male circumcision for HIV prevention in young men in Kisumu, Kenya: a randomised controlled trial. Lancet **2007**; 369:643–56.
14. Gray RH, Kigozi G, Serwadda D, et al. Male circumcision for HIV prevention in men in Rakai, Uganda: a randomised trial. Lancet **2007**; 369:657–66.
15. Fleming DT, Wasserheit JN. From epidemiological synergy to public health policy and practice: the contribution of other sexually transmitted diseases to sexual transmission of HIV infection. Sex Transm Infect **1999**; 75:3–17.
16. Rottingen J-A, William Cameron D, Garnett GP. A systematic review of the epidemiologic interactions between classic sexually transmitted diseases and HIV: how much really is known? Sex Transm Dis **2001**; 28:579–97.
17. Mugo N, Heffron R, Donnell D, et al. Pregnancy is associated with an increased risk for HIV transmission among African HIV-1 serodiscordant couples. Presented at: International Microbicides Conference, Pittsburgh, PA, **2010**.
18. Powers KA, Poole C, Pettifor AE, Cohen MS. Rethinking the heterosexual infectivity of HIV-1: a systematic review and meta-analysis. Lancet Infect Dis **2008**; 8:553–63.
19. Celum C, Wald A, Lingappa JR, et al. Acyclovir and transmission of HIV-1 from persons infected with HIV-1 and HSV-2. N Engl J Med **2010**; 362:427–39.
20. Lingappa J, Thomas K, Hughes J, Baeten J, Fife K, et al. Infected partner's plasma HIV-1 RNA level and the HIV-1 set point of their heterosexual seroconverting partners. Presented at: 18th Conference on Retroviruses and Opportunistic Infections, Boston, MA, **2011**.
21. Campbell MS, Mullins JI, Hughes JP, et al. Viral linkage in HIV-1 seroconverters and their partners in an HIV-1 prevention clinical trial. PLoS one **2011**; 6:e16986.
22. Lee T-H, Sakahara N, Fiebig E, et al. Correlation of HIV-1 RNA levels in plasma and heterosexual transmission of HIV-1 from infected transfusion recipients. J Acquir Immune Defic Syndr **1996**; 12:427–8.
23. Busch MP, Lee LL, Satten GA, et al. Time course of detection of viral and serologic markers preceding human immunodeficiency virus type

1 seroconversion: implications for screening of blood and tissue donors. Transfusion **1995**; 35:91–7.
24. Lavreys L, Baeten JM, Chohan V, et al. Higher set point plasma viral load and more-severe acute HIV type 1 (HIV-1) illness predict mortality among high-risk HIV-1-infected African women. Clin Infect Dis **2006**; 42:1333–9.
25. Pepe MS. Receiver operating characteristic methodology. J Am Stat Assoc **2000**; 95:308–11.
26. Genolini C, Falissard B. KmL: k-means for longitudinal data. Comput Stat **2010**; 25:317–28.
27. Grambsch PM, Therneau TM. Proportional hazards tests and diagnostics based on weighted residuals. Biometrika **1994**; 81:515–26.
28. Fowke KR, Nagelkerke NJD, Kimani J, et al. Resistance to HIV-1 infection among persistently seronegative prostitutes in Nairobi, Kenya. Lancet **1996**; 348:1347–51.
29. Kaul R, Rowland-Jones SL, Kimani J, et al. Late seroconversion in HIV-resistant Nairobi prostitutes despite pre-existing HIV-specific CD8+ responses. J Clin Invest **2001**; 107:341–9.
30. Suy A, Castro P, Nomdedeu M, et al. Immunological profile of heterosexual highly HIV-exposed uninfected individuals: predominant role of CD4 and CD8 T-cell activation. J Infect Dis **2007**; 196:1191–201.
31. Urban TJ, Weintrob AC, Fellay J, et al. CCL3L1 and HIV/AIDS susceptibility. Nat Med **2009**; 15:1110–2.
32. Zhu J, Hladik F, Woodward A, et al. Persistence of HIV-1 receptor-positive cells after HSV-2 reactivation is a potential mechanism for increased HIV-1 acquisition. Nat Med **2009**; 15:886–92.
33. Fox J, White PJ, Weber J, et al. Quantifying sexual exposure to HIV within an HIV-serodiscordant relationship: development of an algorithm. AIDS **2011**; 25:1065–82.