for Public Health Practitioners. Atlanta, Ga: Centers for Disease Control and Prevention;1992.
14. 1987 Michigan Behavioral Risk Factor Survey. *Breast Cancer Screening Knowledge and Behavior*. Lansing, Mich: Michigan Department of Community Health, Center for Health Promotion, Health Surveillance Section; 1988.
15. Conway BJ, Suleiman OH, Reuter FG, Antonsen RG, Slayton RJ. National survey of mammographic facilities in 1985, 1988, and 1992. *Radiology*. 1994;191:323–330.
16. Fintor L, Haenlein M, Fischer R, Udler J. *Mammography: An Analysis of Reimbursement and Quality Assurance Legislation Including Accreditation and Distribution of Units*. Bethesda, Md: National Cancer Institute, Division of Cancer Prevention and Control; 1992.
17. Public Act 56. In: *The Public and Local Acts of the Legislature of the State of Michigan*. Lansing, Mich: Legislative Service Bureau; 1989.

Reporting Sexual Risk Behavior for HIV: A Practical Risk Index and a Method for Improving Risk Indices

Ezra Susser, MD, DrPH, Moise Desvarieux, MD, MPH, and Knut M. Wittkowski, PhD, DSc

ABSTRACT

Objectives. As a means of enhancing public health efforts to control sexual transmission of human immunodeficiency virus (HIV), methods were developed to report on risk behavior in a manner that is comparable and widely interpretable.

Methods. An elementary sexual behavior risk index (the vaginal episode equivalent index) that is in accord with some of the essential knowledge about sexual transmission of HIV is described, and a multivariate ordinal risk (MOR) method that can be used to improve such risk indices is introduced.

Results. An example shows that these approaches are applicable to observational studies of seroconversion.

Conclusions. The MOR represents a powerful new tool to develop valid comparable measures of sexual risk behavior and, thereby, to advance HIV prevention research. (*Am J Public Health*. 1998;88:671–674)

Introduction

In public health efforts to control the HIV epidemic, it is vital to assess the efficacy of strategies designed to reduce sexual transmission of HIV.[1,2] Yet, there is no reliable and accepted way of judging one strategy against another in terms of effects on sexual risk behavior. This paper first describes an elementary sexual behavior risk index that was heuristically defined, and then introduces a method for using the data of observational seroconversion studies to improve upon such risk indices.

The Vaginal Episode Equivalent Risk Index

The prevailing approaches to reporting sexual risk behavior pose dilemmas.[3,4] Focusing on a specific behavior (e.g., episodes of unprotected vaginal sex) evades the most critical question, that is, whether overall sexual risk behavior was affected; a decrease in one behavior may be associated with an increase in another. For instance, decreased vaginal sex may be associated with increased anal sex. If, instead, one examines the proportion of sexual episodes that were unprotected, an awkward paradox may be created: When the total number of episodes decreases, this proportion can increase even though the total number of episodes and risk of HIV transmission decrease. Finally, traditional approaches that classify individuals according to category of risk (e.g., high, medium, or low) tend to be arbitrary and thus vary from one study to another.

Faced with these dilemmas in reporting the results of our own clinical trials[5] and observational studies,[6] we have developed an elementary risk index that is compatible with some of the essential knowledge about sexual transmission of HIV and, specifically, with empirical findings showing anal sex to have higher risk and oral sex to have lower risk than vaginal sex.[7,8] The vaginal episode equivalent (VEE) risk index was defined by the following simple linear function: Risk Score = (number of unprotected vaginal episodes) + (2 × number of unprotected anal episodes) + (0.1 × number of unprotected oral episodes). Risk is reported in units of a vaginal episode equivalent (VEE), equivalent to the risk associated with one episode of vaginal unprotected sex, which is an intuitively meaningful unit. The VEE is eminently practical when data on sexual behaviors and/or HIV status are limited, as they are in the great majority of intervention studies.

The VEE can be refined when the data are extensive. In complex data sets, one might differentiate further, for instance, among types of sex or partners. When the number of anal, vaginal, and oral episodes

Ezra Susser is with the Department of Psychiatry, Columbia University; the New York State Psychiatric Institute; and the Center for Urban Epidemiologic Studies, New York Academy of Medicine, New York City. Moise Desvarieux is with the Division of Epidemiology, Columbia University School of Public Health. Knut M. Wittkowski is with the Department of Medical Biometry, Eberhard-Karls University, Tübingen, Germany, and the HIV Center for Clinical Behavioral Studies, New York City.

Requests for reprints should be sent to Ezra Susser, MD, DrPH, School of Public Health, Columbia University, 722 W 168th St, Box 24, New York, NY 10032.

This paper was accepted September 26, 1997.

TABLE 1—Computation of the Risk Score from a Hypothetical Study of Seven Women (A through G) at Risk for HIV Infection

Woman	Risk Behaviors and Seroconversion Outcome (Anal;Vaginal | Outcome)[a]	Outcome Ranking[b]	A Priori Rankings[c]	Risk Scores[d]
A	(3;1|−)	3.5	7 7 7 7 7 7 7	7.00
B	(3;0|−)	3.5	6 5 6 6 5 6 5	5.60
C	(0;4|−)	3.5	5 6 4 5 6 5 6	5.20
D	(1;0|+)	7.0	4 4 5 3 3 2 2	3.54
E	(0;3|−)	3.5	3 3 3 4 4 4 4	3.46
F	(0;2|−)	3.5	2 2 2 2 2 3 3	2.20
G	(0;0|−)	3.5	1 1 1 1 1 1 1	1.00
		r_s^{*e}	.5 .5 .6 .4 .4 .3 .3	

[a]Number of unprotected anal contacts; number of unprotected vaginal contacts | seroconversion outcome; lines connecting women indicate a priori orderable pairs.
[b]Woman D received the highest outcome rank because she seroconverted. Each of the other women received the average of the other ranks 1–6.
[c]All possible rankings that are in accord with a priori knowledge.
[d]For each woman, the risk score is obtained as the average of her a priori ranks weighted by r_s^*.
[e]Standardized Spearman rank correlation coefficient $r_S^* = (r_S+1)/2$ as a measure of accordance between each of the a priori rankings and the outcome ranking.

has been recorded specifically for each partner—which is not the case in most studies—the index can be extended to the VEE nonlinear; rather than assuming that the additional risk conferred by each unprotected episode with a specific partner is constant, it could be assumed to decline exponentially, as suggested by Bernoulli models.[9]

The VEE is still a relatively arbitrary function, however, and it does not address all of the shortcomings of the prevailing approaches to reporting of sexual risk behavior. In particular, like these other approaches, it rests upon the stringent assumption that the relationship between the frequency (or category) of a sexual risk behavior and the risk of HIV transmission is known (e.g., linear). When the form of the relationship is more complicated, as is likely to be the case,[4–8] the meaning of the results is unclear, especially when sexual episodes with regular and occasional partners are included in the same data analysis. Therefore, methods for improving upon the VEE and other risk indices are needed. We here propose a multivariate ordinal risk (MOR) method for this purpose.

The Multivariate Ordinal Risk Method

The MOR uses both available knowledge about HIV risk factors and observed seroconversion outcomes to determine the components of a risk index. A complete description is beyond the scope of a brief report, but the essential features can be presented by means of a brief example.

For illustration, consider an observational study of the risk of contracting HIV among 7 women (women A, B, C, D, E, F, and G) in monogamous relationships with male partners known to be infected with HIV (Table 1). Data comprise the frequency of 2 risk behaviors, unprotected anal sex and unprotected vaginal sex, and the seroconversion outcomes. This example is constructed to resemble the data from a large observational study.[7]

The MOR begins by computing a risk score for each individual in the study sample based on recently developed methods for analyzing multivariate ordinal data.[10,11] In essence, individuals are ranked in 2 different ways: (1) according to their risk behaviors and our a priori knowledge about the relation of these behaviors to HIV transmission (a priori ranking) and (2) according to their actual seroconversion status (outcome ranking).

Many a priori rankings of the individuals may be compatible with the data on risk behaviors and with a priori knowledge. In our hypothetical example, we assumed a priori that, for each of the 2 risk behaviors, the risk of HIV transmission increases as the frequency of the behavior increases and that 1 anal contact is more risky than 1 vaginal contact. (More detailed a priori rules are allowable; for instance, that 1 anal contact is more risky than 2 vaginal contacts.) Thus, A has the highest and G the lowest risk among all women. A and G always receive the ranks 7 and 1, respectively. The other 5 women (B through F) can be compared pairwise. Some pairs can be ordered; for instance, E is at higher risk than F because E has more vaginal contacts and the same number of anal contacts. On the other hand, B and C are "unordered" in relation to one another, so each may be assigned ranks in either order. Altogether, 7 a priori rankings are compatible with the risk behavior data, corresponding to the 7 ways in which B through F can be ordered.

The actual seroconversions determine the outcome ranking. In our example, only D seroconverted; therefore, D receives the outcome rank 7. A, B, C, E, F, and G receive the same outcome rank of 3.5 $(1+2+3+4+5+6)/6=3.5$.

The risk score for an individual is computed by taking an average of the a priori rankings weighted by the (standardized) correlation between the a priori ranking and the outcome ranking. Thus, the rankings based on prior knowledge are weighted by what actually happened; the a priori ranking that best fits the observed outcomes receives the highest weight. In the example, among the 7 a priori rankings, the third has

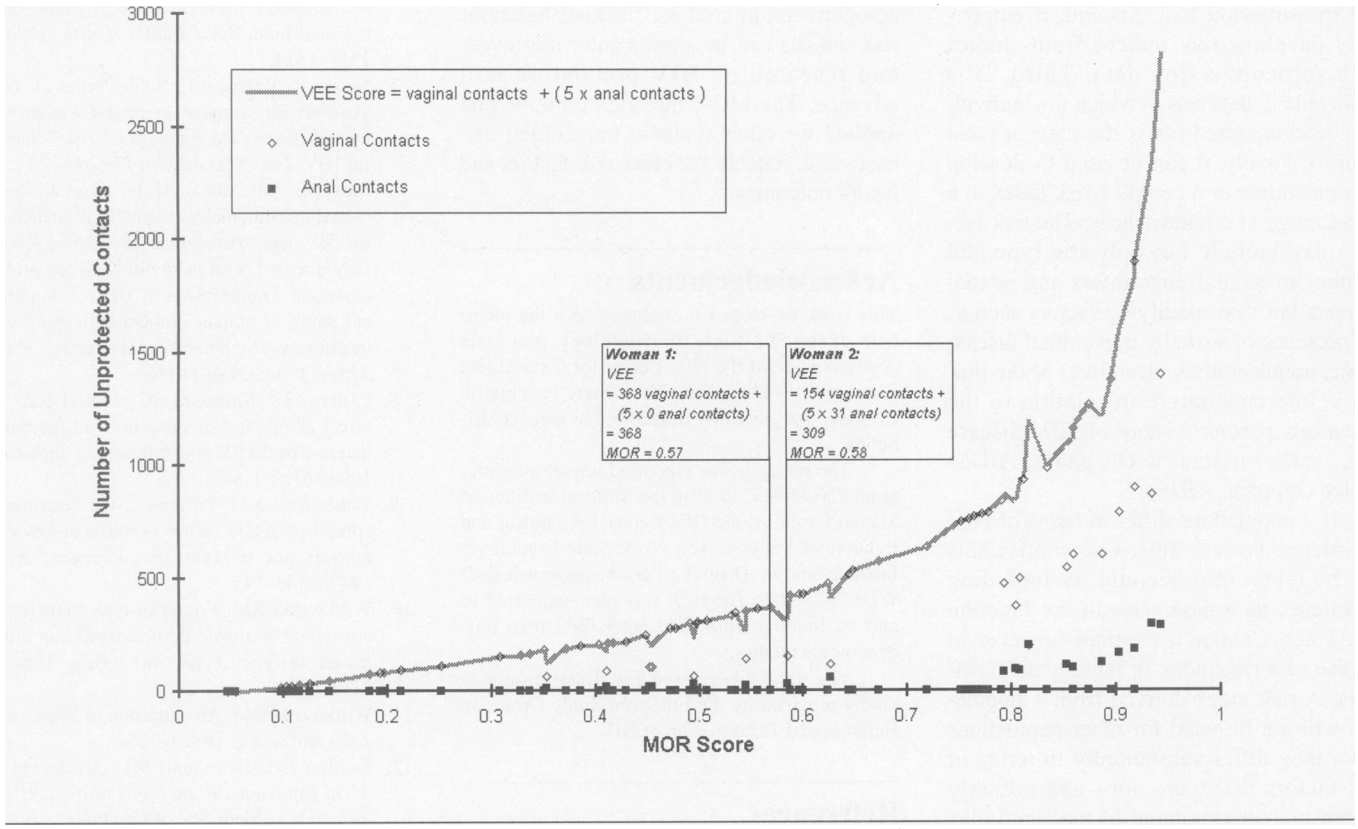

FIGURE 1—Fitting the VEE against the MOR: data from a European study.

the highest correlation (.6) because it assigns the woman who seroconverted (D) a higher rank than the other 6 a priori rankings. The resulting risk score for D is

$$\frac{2\times(5\times4)+(.6\times5)+2\times(.4\times3)+2\times(.3\times2)}{2\times.5+.6+2\times.4+2\times.3}=354$$

Of course, the variables considered by the MOR need not be limited to number and type of sexual episodes, as in this simple illustration. Given sufficient data, we might also have included number of partners or any other risk factors. Further information and software to generate the MOR automatically for large data sets are available on request.

Having obtained the risk scores for the study sample, one now looks for an estimator of the risk for individuals in the population that can be computed when only the risk behaviors are known. This is done by fitting risk scores to a function of the risk behaviors. Such function fitting requires only that risk behaviors and risk scores (obtained by MOR or some other means) are known. In our example, one can fit a linear function reasonably well to the data of Table 1: Risk Score = 1 + (4/5 × number of unprotected vaginal episodes) + (8/5 × number of unprotected anal episodes). This function is similar, but not identical, to the VEE, illustrating that heuristic indices such as the VEE can be refined or validated by the MOR procedure. The choice of a function should be informed by epidemiological models incorporating knowledge of HIV transmission dynamics.[12] As a final step, the estimation procedure may need to be modified so that it can be reported in terms of units or categories that are intuitively meaningful.

Example

To illustrate the use of the MOR to derive risk scores and refine the VEE, we draw on the data of a European study.[7] As in our earlier hypothetical example, the study included women in monogamous relationships with men infected with HIV. Two risk behaviors were considered: unprotected vaginal sex and unprotected anal sex. The MOR procedure was used to generate risk scores as described earlier to a total of 291 women with 56 seroconversions. Then we fitted a linear function—VEE = vaginal contacts + (weight × anal contacts)—to the MOR scores of a subsample of 77 women in couples with no history of other sexually transmitted diseases. As shown in Figure 1, we obtained a "least square" fit when setting the weight to 5.

Discussion

This paper has introduced an elementary (VEE) sexual risk index that is in accord with essential knowledge about HIV and that can be used in the majority of HIV intervention studies in which data are limited. In addition, we have proposed a method—termed MOR—for using the data of observational studies of HIV seroconversion to refine and validate risk behavior indices such as the VEE or to derive indices de novo.[13] In efforts to define valid comparable measures of sexual risk behaviors, the MOR offers a powerful new tool that can be adapted to complex situations.

The MOR method has a number of desirable features. First, it can be used with

minimal a priori assumptions about the nature of the relationship between behavior and transmission risk. Second, it empirically develops risk indices from studies with seroconversion data. Third, it is applicable to data sets in which few individuals seroconverted (as is the case in most studies). Finally, it can be used to develop either a simple or a complex risk index in a broad range of circumstances. The risk factors may include not only the type and number of sexual encounters and sexual partners but also modifying factors such as the presence of sexually transmitted disease (none, nonulcerative, ulcerative) or the timing of the encounters in relation to the infectious partner's stage of HIV disease (e.g., early viremia, latent phase, AIDS-related complex, AIDS).[14]

If 2 populations differ in terms of HIV prevalence but are otherwise similar, this can be taken into account by including prevalence as a parameter in the function fitting step. Caution is required, however, in the use of a risk index in more general settings. A risk index derived from 1 population will not be valid for other populations when they differ substantially in terms of risk factors that were not—and possibly cannot be—incorporated in the risk index or in terms of other factors such as the predominant strain of HIV. Therefore, the elemental frequencies of behaviors that compose a risk index (e.g., number of episodes of vaginal and anal sex) should be routinely reported.

Conclusion

The VEE risk index provides a ready means to improve reporting of HIV sexual risk behavior data in both observational and intervention studies. This index, however, still has significant shortcomings. By applying the proposed MOR method to data in seroconversion studies,[13] sexual behavior risk indices can be significantly improved, and research on HIV prevention will advance. The MOR method could also be applied for other sexually transmitted diseases and, indeed, for other risk factors and health outcomes.[15] □

Acknowledgements

This work developed in dialogue with the members of the Statistics, Epidemiology, and Data Analysis Core of the HIV Center for Clinical and Behavioral Studies, New York State Psychiatric Institute; we gratefully acknowledge their contribution.

The research was supported in part by center grant P50-MH43520 from the National Institute of Mental Health to the HIV Center for Clinical and Behavioral Studies, New York State Psychiatric Institute, and by Deutsche Forschungsgemeinshaft WI 827/7-1 Dfg Grant. It was also supported in part by funding from the Center for Urban Epidemiologic Studies.

The data subset used for illustration was kindly provided by the European Study Group on Heterosexual Transmission of HIV.

References

1. Catania JA, Coates TJ, Stall R, et al. Prevalence of AIDS-related risk factors and condom use in the United States. *Science.* 1992;258:1101–1106.
2. Cates W Jr. Contraception, unintended pregnancies and sexually transmitted diseases: why isn't a simple solution possible? *Am J Epidemiol.* 1996;143:311–318.
3. Samuel MC, Mohr MS, Speed TP, Winkelstein W Jr. Infectivity of HIV by anal and oral intercourse among homosexual men. Estimates from a prospective study in San Francisco. In: Kaplan EH, Brandau ML, eds. *Modeling the AIDS Epidemic: Planning, Policy, and Prediction.* New York, NY: Raven Press Ltd; 1994.
4. Joseph JG, Adib SM, Koopman JS, Ostrow DG. Behavioral change in longitudinal studies: adoption of condom use by homosexual/bisexual men. *Am J Public Health.* 1990;80:1513–1514.
5. Susser E, Valencia E, Sohler N, et al. Interventions for homeless men and women with mental illness: reducing sexual risk behaviors for HIV. *Int J STD AIDS.* 1996;7:66–70
6. Susser E, Valencia E, Miller M, et al. Sexual behaviors of homeless mentally ill men at risk for HIV. *Am J Psychiatry.* 1995;152:583–587.
7. DeVincenzi I, European Study Group on Heterosexual Transmission of HIV. A longitudinal study of human immunodeficiency virus transmission by heterosexual partners. *N Engl J Med.* 1994;331:341–346.
8. Padian NS, Shiboski SC, Jewell NP. The effect of number of exposures on the risk of heterosexual HIV transmission. *J Infect Dis.* 1990;161:883–887.
9. Wittkowski KM. Preventing the heterosexual spread of AIDS: what is the best advice if compliance is taken into account? *AIDS.* 1989;3:143–145.
10. Wittkowski KM. Friedman-type statistics and consistent multiple comparisons for unbalanced designs. *J Am Stat Assoc.* 1988;83:1163–1170.
11. Wittkowski KM. An extension to Wittkowski. *J Am Stat Assoc.* 1992;87:258.
12. Kaplan EH, Brandeau ML. *Modeling the AIDS Epidemic: Planning, Policy, and Prediction.* New York, NY: Raven Press Ltd; 1994.
13. Desvarieux M, Wittkowski KM, DeVincenzi I, Susser E. A new objective method for computing risk indices applied to cohorts of HIV-discordant couples. In: International Conference on AIDS, July 1996, Vancouver, British Columbia, Canada. Abstract MOC1608.
14. Jacquez JA, Koopman JS, Simon CP, Longini IM Jr. Role of the primary infection in epidemics of HIV infection in gay cohorts. *J AIDS.* 1994;7:1169–1184.
15. Einsele H, Ehninger, G, Hebart H, et al. Polymerase chain reaction monitoring reduces the incidence of cytomegalovirus disease and the duration and side effects of antiviral therapy after bone marrow transplantation. *Blood.* 1995;86:2815–2820.