RESEARCH REPORT

Self-reports of HIV risk behavior by injecting drug users: are they reliable?

MARJORIE F. GOLDSTEIN, SAMUEL R. FRIEDMAN, ALAN NEAIGUS, BENNY JOSE, GILBERT ILDEFONSO & RICHARD CURTIS

National Development and Research Institutes, Inc., 11 Beach Street, New York, New York 10013, USA

Abstract
While most studies of AIDS risk behavior rely on self-reports, few studies have assessed the reliability of these reports. The present study examines self-reports of drug-related and sexual risk behavior among pairs of injecting drug users (IDUs) recruited from the streets in New York City. Since both members of the pair were interviewed, it was possible to compare their responses in order to assess reliability. Subjects reported on their contacts' demographic data (age, gender, race/ethnicity) and on shared risk behaviors, including syringe sharing. Despite the private and/or illegal nature of AIDS risk behaviors, IDU subjects were generally reliable in their reports of both demographic and AIDS risk behaviors.

Introduction
Most AIDS/HIV risk research depends on self-reports of sexual and drug injecting behavior. In many studies of HIV epidemiology and prevention, the behavior of injecting drug users (IDUs) is the main focus. Although it has often been assumed that IDUs were unreliable reporters of their own behavior, this is no longer accepted without question. In this paper we report on the reliability of data provided by a sample of street-recruited IDUs on their own AIDS/HIV transmission risk behaviors. (Amsel et al., 1976; Ball, 1967; Bonito, Nurco & Schaffer, 1976; Maddux & Desmond, 1975; Stephens, 1972). The assumption that morphine (and later heroin) addicts in particular, are liars comes from unsupported statements in the medical literature, which were first made in the last century but which continued to be published well into this century in at least one standard medical text (Osler, 1892; and Christian, 1944 cited in Maddux & Desmond, 1975). Harrell (1985) points out that most studies of the 'verdictality' of self-reported data from current or former drug users failed to find denial or concealing of drug use. Where empirical evidence supporting (or refuting) the assumption was available, it was largely limited to IDUs who were in treatment or incarcerated, i.e. to people who had something to lose or gain from truthful reporting. The information collected from these

Background
Evidence that IDUs are "untruthful" is scanty

Correspondence to: Marjorie F. Goldstein, National Development and Research Institutes, Inc., 11 Beach Street, New York, New York 10013, USA.

0965-2140/95/081097-08

© 1995, Society for the Study of Addiction to Alcohol and Other Drugs
subjects mainly concerned general drug use, arrests or hospitalizations. Prior to the advent of AIDS, there was little interest in either the sexual behavior of IDUs or in the sharing of drug paraphernalia as a means of disease transmission.

The lack of data from earlier research highlights the influence of the time period in which the studies were conducted. The pre-AIDS period (roughly prior to 1981) can be contrasted with the post-1981 AIDS era. Notably, there has been a shift in interest about the types of IDU behavior under investigation from the first era to the second: that is, from measuring illegal activities and specific drugs used, to measuring drug and syringe-sharing practices and sexual behaviors. Findings from research in each era, while ostensibly covering some of the same topics—i.e., drug use—are not comparable in many ways.

For example, in the pre-AIDS era, most studies were conducted on in-treatment (or incarcerated) populations, where drug use was a criterion of both personal and programmatic success (or failure). For an individual study subject, drug use of a certain level often meant dismissal from the program or withholding of privileges. Route of drug administration (needle vs. other) was not of great concern. After urine testing was developed, measurement of the time period of drug use was reduced. That is, self-report measures were usually taken for the very limited period (approximately 48 hours prior to testing) for which biological verification was available (Mieczkowski et al., 1991; Wish et al., 1983).

By contrast, in the AIDS era much research on drug use has focused on those behaviors thought to be most related to AIDS/HIV transmission. These include: duration of drug use, details of injection procedures, especially the sharing of drug injecting equipment, frequency of drug use and location of drug use, including at shooting galleries. In addition, since HIV is known to be sexually as well as parenterally transmissible, there has been considerable interest in the sexual practices of IDUs. These include: condom use, number of partners, types of partners, types of sexual activities and frequency of sexual activity.

The type of IDU population of interest to researchers has also changed. While in the pre-AIDS era most of the IDU populations studied were in drug treatment or prison, in the AIDS era street populations as well as in-treatment and prison populations are studied.

There has also been a shift in attitude on the part of some researchers toward their IDU subjects. Drug use has been less stigmatized by these investigators. That is, while some researchers continue to view drug use as behavior very negatively, others have adopted a more neutral stance. This stance includes the "harm reduction philosophy" which views drug use as an adaptation some people exhibit towards a harsh environment (Springer, 1991). From a harm reduction approach, researchers not only study how IDUs who wish to stop using drugs can be helped to do so, but also how other IDUs can be helped in staying as healthy as possible (Des Jarlais, Friedman & Ward, 1993). This is at odds with the criminal justice system, which still regards most drug use as a crime, and even with much of the medical system which considers drug use to be an avoidable illness. Those who regard drug use more neutrally have been more likely to question the assumption that drug users are "untruthful" by their very nature. This questioning has spurred research on self-reports of behavior relevant to HIV transmission.

Studies on the reliability of IDUs' self-reports are now conducted similarly to other studies of self-reports, namely by comparing self-reports to others' reports or to written records (Darke et al., 1991; Padian, 1990). The concept of reliability has a number of different definitions (Steglich, Wrightman & Cook, 1976), including repeatability, internal consistency and agreement among observers. In the study to be described reliability is operationalized in several ways, depending on the type of information being considered and its most accurate source. Regarding demographic and historical data, reliability is defined as how well a second person's report agrees with a subject's own report. When considering shared or mutual behavior, reliability is assessed via agreement between the two persons, since neither is considered to be inherently more accurate than the other.

Because the behaviors under study are important to the spread of HIV, their modification could reduce that spread. It is virtually impossible to assess these behaviors other than through self-reports. This study afforded a unique opportunity to assess reliability of such self-reports by comparing a self-report to that of another observer, namely another IDU who sometimes injected with, and perhaps had sex with, the subject.
Study methodology

Sampling and subject linkage

The "Social Factors and HIV Risk" study utilized a chain referral sampling technique in order to study IDU social networks. Using this method, IDUs were recruited in a neighborhood with high rates of drug use and high seroprevalence of HIV among IDUs. They were interviewed and asked to nominate people in their social networks with whom they had contact in the past 30 days. Subjects were screened for eligibility (at least 18 years old, having injected drugs in the past year), scheduled for an interview and offered an HIV test, with informed consent and appropriate pre- and post-test counseling. (A more complete discussion of sampling and linkage procedures can be found in Neaigus et al., in press.)

In the interview, subjects were asked to provide detailed sexual and drug histories. They were also asked specific questions about their injection and sexual behaviors in the past 30 days, if any, with each of their nominated network members. Attempts were then made to recruit their IDU nominees for an interview and HIV test.

After one or both of two IDUs had nominated the other as a network member (contact), this linkage was validated in one or more of four ways: (1) when a previously nominated network member arrived at the research storefront accompanied by the index subject, the pair were considered to be linked; (2) through a field link in which the subject pointed out a nominated contact to a research staff member; (3) through an ethnographic link in which two subjects were seen together in a social or drug-related interaction by project staff members, who then performed a data base check which confirmed that at least one of the two subjects had nominated the other; or (4) in the study's field data base, through matching on a number of key descriptive variables (age within 5 years, first name or street name, height within 3 inches, ethnic group), again checked by confirming nominee status (see Neaigus et al., in press).

Links were then categorized as unidirectional or bidirectional. In order to be a link at all, at least one member of the pair had to nominate the other, thus forming a unidirectional link. If each person nominated the other, a two-way link was formed.

The linkage process highlights the key role of ethnography as a research tool in this study. Without the work of the ethnographic team it would not have been possible to recruit so many pairs of drug injection partners, nor would it have been possible to make field, ethnographic or data base links as described above.

Sample characteristics

The subjects were recruited during the period July 1991 to January 1993, as described above, in the Bushwick section of Brooklyn in New York City. Most residents of the area are Latino, although Bushwick borders on other neighborhoods, some predominantly black and others predominantly white. People who buy or use drugs in the neighborhood are a mixture of these racial/ethnic groups. The final study sample of subjects for whom linked data are available consisted of 423 persons who entered into one or more links. (From these 423 individuals, 521 unique pairs were constituted; 147 of which had bidirectional links and 374 of which had only a unidirectional link.) These subjects were 28% black, 35% Latino, 36% white and 1% other. Their mean age was 34.8 years (range: 18–64 years). The sample was 69% male and 31% female. Twenty per cent of subjects said they were homeless at the time of the interview and 53% had incomes of less than $10,000 per year, indicating that the sample was primarily poor. Subjects had been injecting for an average of 14 years (range: <1 year–44 years). Twenty-two per cent of subjects said that they were in drug treatment at the time of the interview. The overall sample HIV seroprevalence among those tested (89% of sample) was 41% positive.

Criteria for assessing reliability

Reliability was assessed in three different ways, depending on the type of question asked and the type of data (continuous, categorical or binary). For continuous variables (e.g., age), Pearson correlation coefficients were calculated in order to ascertain the degree of association. In order to further assess reliability, paired t-tests were performed on the responses provided by subjects and contacts for that variable. Significance tests allow for examination of the null hypothesis that there is no difference between the two responses. Where statistically significant differences were found, the magnitude of the difference in means
allowed us to assess the seriousness of the discrepancy. For categorical variables (e.g., race/ethnicity), percent agreement was calculated. For agreement on whether or not some mutual action has taken place (binary), both percentage agreement and Kappa (Fleiss, 1981) were calculated. Significance of Kappa was evaluated using the asymptotic standard error (Norusis, 1992).

**Analyses**

**Demographic characteristics.** The purpose of this analysis is to determine how well subjects are able to report on the age, gender and race/ethnicity of their nominated contacts. The standard used to judge “correctness” is the report of the nominated contact him/herself.

**Years of injection.** This is a critical historical variable in a study of IDUs since it has been shown to be significantly related to HIV status (Schoenbaum et al., 1989). Again, the standard is information from the nominated contact him/herself. A subanalysis was performed for this variable in which respondents’ answers were stratified by length of time they reported knowing the identified contact.

**Shared behavior.** This is another important set of variables for studying HIV transmission as well as for studying social interactions. As described above, only two-way linkages could be used for this type of analysis. It should be noted that there is no “correct” answer by which to judge either subject’s answer. Therefore, both percentage agreement and Kappa were used to evaluate agreement between the two subjects for binary variables, and Pearson correlation coefficients were calculated for continuous variables.

First, agreement on specific behaviors was considered. Data on two types of behavior were analysed, “ever injected drugs with this contact” and “ever had sex with this contact”. Next, agreement was assessed on the duration of the relationship by type, either drug injection or sexual. Then, if the relationship was sexual, agreement on condom use was assessed. Finally, if the relationship included drug injecting, agreement was assessed on two very specific behaviors in the past 30 days: “any lending of a syringe to this contact”, and “any borrowing of a syringe from this contact”. These latter two behaviors are actually complementary halves of the same exchange. Pairs of individuals were, therefore, evaluated for agreement on “lending—borrowing” and “borrowing—lending” separately. The results were then summed in order to obtain percentage agreement on whether the complementary halves were mentioned by each member of the pair.

**Observed behavior.** In this case the subject’s observation of a contact’s behavior is evaluated against what that contact states about his/her own behavior. For purposes of this analysis, these behaviors are dichotomized into “ever” and “never” to facilitate calculation and because the observer and direct respondent may have had different base rates upon which they estimated the frequency of particular behaviors (going to shooting galleries, for example). Percentage agreement is calculated and, although the two observers are not exactly equivalent, Kappa is also calculated.

**Results**

**Demographic data**

Subjects reported on the age, sex, and race/ethnicity of their contacts. The Pearson correlation coefficient for subject’s estimation of contact’s age by that contact’s self-reported age is 0.79 ($p < 0.0001$) (Table 1). There was 100% agreement between respondents and their contacts on the gender of the contact and 92% agreement on race/ethnicity. Percentage agreement on race/ethnicity was calculated by adding agree-agree responses for four categories (black, Latino, white and other) and dividing by all responses.

**Contacts’ years of injection (see Table 1)**

Subjects were asked to estimate their contacts’ years of injection via the question “How long would you guess that this person has been injecting drugs?” The correlation coefficient with the contact’s own self-report data was 0.56 ($p < 0.0001$). In order to determine whether those of longer acquaintance would be more reliable, the sample was stratified by how long the subject had known the contact. For those who knew the contact more than 5 years, the correlation between contact’s actual injection years and the subject’s “guess” was 0.70 ($p < 0.0001$), while for those who had known the contact for 5 years or less, the correlation
Reliability of IDU risk behavior self-reports

Table 1. Pearson correlations: subjects’ and contacts’ responses

<table>
<thead>
<tr>
<th>Variable</th>
<th>$r$</th>
<th>$n$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>0.79</td>
<td>642</td>
<td>0.0001</td>
</tr>
<tr>
<td>Years of injection (contact’s $n$ of years vs. subject’s “guess”)</td>
<td>0.56</td>
<td>551</td>
<td>0.0001</td>
</tr>
<tr>
<td>If subject knows contact 5 years or less years of injection</td>
<td>0.40</td>
<td>306</td>
<td>0.0001</td>
</tr>
<tr>
<td>If subject knows contact more than 5 years years of injection</td>
<td>0.70</td>
<td>245</td>
<td>0.0001</td>
</tr>
<tr>
<td>Duration of injecting relationship (in months)</td>
<td>0.71</td>
<td>124</td>
<td>0.0001</td>
</tr>
<tr>
<td>Duration of sexual relationship (in months)</td>
<td>0.89</td>
<td>49</td>
<td>0.0001</td>
</tr>
</tbody>
</table>

$n =$ number of links used in each analysis. Age is based on all links; years of injection is based on all links where data are available; duration of injecting or sexual relationship is based on bidirectional links where both IDUs agreed that they had such contact during the past 30 days.

The correlation coefficient was only 0.40 ($p < 0.0001$). While both of these latter estimates of age showed a statistically significant difference between the subjects’ reports and the contacts’ guesses (paired $t$-test $p < 0.001$), the mean differences were only 2.9 (95% CI, 2.0–3.9) years, and 2.8 (95% CI, 2.0–3.7) years, respectively. This indicated that subjects tended to underestimate the length of time their contacts had been injecting drugs.

Shared behaviors (Table 2)

Subjects reported on several shared behaviors. There was 93% agreement on ever having injected drugs with each other (Kappa = 0.34) and 99% on ever having had sex with each other (Kappa = 0.97). Regarding the duration of drug injecting relationships, the Pearson correlation coefficient between members’ reports (in months) was 0.71 ($p < 0.0001$; paired $t$-test for difference between reports, $p = 0.35$); for the duration of sexual relationships it was 0.89 ($p < 0.0001$; paired $t$-test for difference between reports, $p = 0.76$). In reporting on condom use, agreement was 82% (Kappa = 0.41). When reporting on having borrowed or lent a syringe to the other during the past 30 days, agreement was 71% (Kappa = 0.39).

Observed behaviors (Table 2)

Two behaviors known to be related to risk of HIV infection are use of shooting galleries and other types of multi-user setting (Chaisson et al., 1987; Jose et al., 1993; Marmor et al., 1987, Neaigus et al., 1992; Schoenbaum et al., 1989). Data were collected in detail on both the shooting galleries and the other multi-user settings where a subject had injected in the past 30 days. Subjects were also asked where his or her contact(s) had injected in the past 30 days. Responses were dichotomized into “yes, in past 30 days” and “no, in past 30 days” for reports on both index subjects’ and contacts’ behaviors. There was 74% agreement (Kappa = 0.32) among 474 pairs as to whether a person had injected in a shooting gallery, and 71% agreement (Kappa = 0.25) among 503 pairs as to whether a person had injected in another type of multi-user setting in the past 30 days.

Discussion

Results of the analyses show that IDUs in this sample are reliable reporters of their own HIV/AIDS risk behaviors as well as of other important information (years injecting, demographics). The type of reliability assessed here was a form of inter-observer or inter-rater reliability. While test-retest reliability is also of interest for HIV/AIDS risk behavior, it could not be estimated from these data. The results, however, contradicted the assumption that ‘addicts are liars’ and that one cannot rely on information they provide.

Two important aspects of the study may be related to the high reliability obtained. One is the complete separation of the study, both organizationally and philosophically, from law enforce-
Table 2. Agreement between members of linked pairs on shared actions

<table>
<thead>
<tr>
<th>Dichotomous variables</th>
<th>Percentage agreement</th>
<th>Kappa</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ever injected together (life-time)</td>
<td>93</td>
<td>0.34</td>
<td>0.01</td>
</tr>
<tr>
<td>Ever had sex together (life-time)</td>
<td>99</td>
<td>0.97</td>
<td>0.001</td>
</tr>
<tr>
<td>Condom use (ever/never in past 30 days)</td>
<td>82</td>
<td>0.41</td>
<td>0.01</td>
</tr>
<tr>
<td>Lent/borrowed each other’s syringe in past 30 days</td>
<td>71</td>
<td>0.39</td>
<td>0.01</td>
</tr>
<tr>
<td>Hang out together (ever/never) in past 30 days</td>
<td>72</td>
<td>0.35</td>
<td>0.001</td>
</tr>
</tbody>
</table>

ment and drug treatment. This suggests that, in studies conducted at criminal justice or drug treatment agencies, special care needs to be taken to reduce the probability that responses might be influenced by fear of potential penalties associated with the behaviors about which subjects are being interviewed. The second is the recruitment and selection of subjects without regard to their treatment status. Although data were gathered on past and present drug treatment status, no attempts were made to urge subjects to enter or return to treatment. (Referrals to various treatment programs were made to subjects after the interview, however, upon request.) These two factors were likely to have contributed to an atmosphere of trust and to the subjects’ likelihood of providing honest answers, since they clearly had nothing to lose or gain by misrepresenting themselves. In studies of untreated or incarcerated IDUs, there is often something at stake for the individual subject.

Another factor which probably contributed to the accuracy of the self-reports is the generally neutral stance on drug use subscribed to by the project staff. While not all members of the study team would adhere exactly to the “harm reduction” philosophy referred to earlier, all staff believed that subjects were people worthy of consideration and respect who had valuable information to provide regarding HIV risk.

Although this is not the only study to have these characteristics, it is one of the few to estimate subjects’ reliability, especially with respect to shared behavior. The calculation of agreement on “lending—borrowing” and “borrowing—lending” of syringes by members of drug sharing pairs may be unique.

The possibility exists that a change in attitude has also occurred in the IDU population during the AIDS era. That is, most IDUs are aware of the connection between injecting drug use and risk of HIV infection (although they may or may not have accurate harm reduction information). IDUs are also aware of researchers’ interest in them. These two factors may motivate IDU research subjects to be more truthful than in the pre-AIDS era.

Two limitations of this study are the non-random nature of the sample and the lack of two-way data for many of the links (which, in turn, limits the sample size for analyses of shared behavior). These limit the generalizability of the findings. Generalizability cannot be established with regard to other groups until this type of data is gathered in other IDU populations. To do this, smaller reliability studies of pairs of IDUs might be profitably added to many studies.

A third limitation is the absence of validity measures, that is, objective measures of the behaviors under investigation. Currently no truly objective measures are available for the behaviors of interest in this study. While urine tests are available to measure use of specific drugs during the 48 hours prior to the test (Wish et al., 1983), drug use with others in the past 30 days was the risk behavior under consideration here. It is difficult to devise measures of validity for shared risk behaviors (Catania et al., 1990). Some investigators consider corroboration by the other member of a pair to be a measure of validity (Darke et al., 1991; Kleyn, Day & Weis, 1991).

The generally low range of Kappa values presented deserves some comment. A low Kappa can be either the result of poor percentage agreement or a skewed distribution of cases among the cells of the tables (Pleiss, 1981; Landis & Koch, 1977; Zwick, 1988). Since such skewed distributions did occur in the data, some Kappa
values were low, even when percentage agreement was high. The Kappa values reported in this paper are all statistically significant, however, indicating greater than chance agreement by the two IDU raters. The sample for this study, while not randomly selected, included subjects of both genders, members of the predominant racial/ethnic groups in New York City, both in-treatment and out-of-treatment subjects and subjects with a wide range of both age and years of drug injection. The sample was, therefore, at least broadly representative of IDUs in New York City.

Conclusion
Subject to the limitations discussed above, this study demonstrates that IDUs’ self-reports about their injecting and sexual behaviors are generally accurate about behaviors which are related to HIV infection. IDU subjects appear to be reliable informants even about their illicit drug use, including injection procedures, and their private sexual behaviors. The findings are encouraging for the field of HIV/AIDS risk research. Researchers who are inclined to distrust self-reports by IDUs can miss the opportunity to gain information unobtainable by any other means.

Acknowledgements
Support was provided by USA National Institute on Drug Abuse grant #RO1 DA06723. The New York State Department of Health AIDS Institute assisted in drawing serum samples, HIV testing, and referrals.

References


