Validity of Drug Use Reporting in a High-Risk Community Sample: A Comparison of Cocaine and Heroin Survey Reports with Hair Tests

Michael Fendrich,1 Timothy P. Johnson,2 Seymour Sudman,3 Joseph S. Wislar,1 and Vina Spiehler4

Hair specimens were collected from 322 subjects and analyzed as part of an experimental study administering household surveys during 1997 to a high-risk community sample of adults from Chicago, Illinois. Toxicologic results were compared with survey responses about recent and lifetime drug use. About 35% of the sample tested positive for cocaine, and 4% tested positive for heroin. Sample prevalence estimates of cocaine use based on toxicologic results were nearly five times the survey-based estimates of past month use and nearly four times the survey-based estimates of past year use. With the hair test results as the standard, cocaine and heroin use were considerably underreported in the survey. Underreporting was more of a problem for cocaine than for heroin. Among those who tested positive, survey disclosure of cocaine use was associated with higher levels of cocaine detected in hair. In general, when recent drug use was reported, it was usually detected in hair. When a drug was detected in hair, use was usually not reported in the survey. When heroin was detected in hair, cocaine was almost always detected as well. Am J Epidemiol 1999;149:955-62.

Increasing concerns have been raised about the validity of drug use prevalence estimates based on survey reports (1–3). Toxicologic advancements, particularly those that facilitate the detection of cocaine and opiates in hair (4–6), promise to enhance our understanding of the limits of drug use surveys. Studies incorporating both hair testing and survey research methods have used the standard of the hair test to evaluate the accuracy of reports of recent drug use (7–10). To date, most studies that use these comparative procedures have focused on specialized samples in limited contexts, such as inmates in jail or prison (11–14), addicts in drug treatment programs (7), obstetric patients (9), or employees on the job (10). Since national drug use prevalence estimates are most commonly based on results from household surveys conducted in the community, insights concerning the limitations of drug use reporting in this context are of critical importance (15). This report summarizes the results of one of the first studies to directly compare hair test results with survey responses from household interviews obtained from a community sample.

MATERIALS AND METHODS

Sample and data collection

The main study from which this sample is derived was based on a community sample of residents aged 18–35 years who resided within the city of Chicago, Illinois, in 1997. The pool of neighborhoods selected for the study was restricted to those with above-average admissions to state-supported drug and alcohol treatment programs. In addition, neighborhoods were stratified according to 1990 Census data to assure a balanced distribution across race/ethnicity groups (i.e., one third of the neighborhoods were predominantly African American, one third were predominantly white, and one third were predominantly Hispanic). Randomly selected blocks from eligible neighborhoods were randomly assigned to one of two interview conditions—a "control" condition, closely following the procedures and instruments used in the 1995 National Household Survey on Drug Abuse (16) and an experimental condition that incorporated cognitive interview procedures (17). The latter condition included
the use of a computer-assisted, self-administered interview, clearer definitions for the drugs inquired about, and extensive cognitive probes designed to stimulate and facilitate recall of recent behavior. Within selected households, a variant of the Troldahl-Carter-Bryant procedure (18, 19) was used to select randomly from among eligible respondents aged 18–35 years. During the screening phase of the interview, informants from selected households were asked to participate in an interview “for our study on alcohol, tobacco, and other substances.” Forty-seven percent of eligible informants agreed to participate.

Trained interviewers oriented each respondent to the assigned protocol. Each household interview protocol included detailed questions about past drug use of a total of 12 different substances. Questions were also asked about impairment resulting from drug use and recent involvement in drug treatment. Respondents were not offered a financial incentive for participation in either of the interview conditions. Upon the conclusion of the interview, however, interviewers asked respondents to provide a specimen of their hair for a $10 incentive. As part of the consent process, respondents were told that hair samples would be tested for drugs. If a respondent consented to participate, interviewers cut approximately 60 strands of hair from the head. For three cases, samples were taken from facial hair. Respondents were not told about the hair test until after they had completed the drug use survey. Hair segments were measured, wrapped in foil with the root ends marked, sealed in a paper envelope, and enclosed in a plastic bag. Specimens were sent to United States Drug Testing Laboratories in Des Plaines, Illinois, for toxicologic analysis.

Variables

**Drug use reporting.** Three sets of parallel drug use reporting measures were derived from the two protocols. Respondents were asked about their lifetime and most recent use of cocaine and/or crack (labeled in the remainder of this report as “cocaine”) and heroin. On the basis of responses to these questions, subjects were classified as lifetime, past year, past month, and nonusers of each of these substances.

**Passive exposure.** Researchers have identified the potential for passive exposure to drugs, especially cocaine powder, to result in positive hair tests results (20). To adjust for the potential influence of passive exposure on hair test results, all respondents were asked whether they had recently been in the same room with someone ingesting cocaine. Subjects who reported that they had been in the same room with someone ingesting this substance on at least 1 day during the past 30 days were counted as having experienced passive exposure.

**Hair cocaine and opiate analysis.** Hair segments were washed and then screened using a fluorescence polarization assay process as previously described (21). Analyses were performed on hair segments cut to 3 cm in length from the root ends, corresponding to about 3 months hair growth (22). Samples shorter than 3 cm were analyzed in toto. All specimens exceeding 0.05 ng/mg of cocaine or its metabolic equivalents (benzoylcegonine or cocaethylene) were confirmed for cocaine using a gas chromatography/mass spectrometry procedure. Similarly, all specimens exceeding 0.5 ng/mg of opiate equivalents (codeine, morphine, or monoacetyl morphine) were confirmed for opiates using gas chromatography/mass spectrometry. All confirmed cocaine samples that identified cocaine or its metabolites at a level exceeding 0.01 ng/mg were considered to be cocaine-positive specimens. All confirmed opiate samples that identified opiates or their metabolites at a level exceeding 0.05 ng/mg were considered to be opiate-positive specimens. Toxicologic analyses of hair suggest that monoacetyl morphine or morphine are most commonly detected in heroin users (6). Accordingly, only confirmed specimens that identified either of these metabolites (either alone or in combination with codeine) were considered positive tests.

**Background variables.** Since prior research (23–25) suggests that black informants may be less willing to disclose drug use than are white informants, race/ethnicity comparisons are of particular interest. Other background variables examined as potential correlates of drug use disclosure were gender, highest level of education completed, employment status, age, and whether a subject was assigned to the control or experimental condition of the interview.

**Statistical analyses**

Unweighted sample prevalence estimates were compared across measures within each substance. Paired differences between the proportion indicating use on the survey and that testing positive were evaluated using the McNemar test (26). Agreement between drug tests and interviewer reports was evaluated by examining sample estimates for sensitivity (using the hair test as the standard for a positive result), coefficient kappa (26), and conditional kappa (27). Conditional kappa is an appropriate statistic when one of the indicators (or ratings) in a comparison is considered to be a “standard.” Since biologic measures are often considered the standard in drug use studies, conditional kappa has been used to evaluate agreement between biologic measures and survey reports (7, 8, 24). Thus, in our study, the conditional kappa evaluated chance-corrected agreement only for those subjects.
who had been classified by the hair test as “positive” for a particular drug (i.e., agreement level conditional on a positive drug test). Since prior research suggested that drug users were more comfortable disclosing use in more remote periods of time (24), agreement between drug tests and lifetime, past year, and past month reports were evaluated. Analyses also evaluated the sensitivity and specificity of interview reports for each of these time periods. Bivariate cross-tabular analyses evaluated the association between interview condition and background variables and drug use disclosure among those with positive cocaine hair tests.

RESULTS

Hair specimens were obtained from 322 respondents, or about 56 percent of the sample of 568 respondents. Compared with the rest of the sample, those with usable specimens were significantly more likely to be female, white or other, and unemployed (table 1).

As a verification of the sensitivity of the laboratory procedures, we split three samples known to be positive for opiates and cocaine by prior toxicologic analyses and submitted them as blind specimens to the laboratory on two occasions. Two of the specimens were obtained by the study toxicologist (V. S.); these specimens originated in Germany, where they were used as part of an interlaboratory round-robin proficiency test. The third specimen was reference material supplied by the National Institute of Standards and Technology. The laboratory confirmed cocaine or its metabolites in all six specimens. Opiates were identified in only four of the six specimens, however. As a check on laboratory specificity, 10 known negative samples (specimens provided by the investigators and their associates) were sent to the laboratory during the course of the study. All 10 of these negative samples tested negative.

Results of comparisons of sample prevalence rates across each of the measures evaluated in the study are presented in table 2. Note that 111 specimens were positive for cocaine, and 13 were positive for heroin. Eleven of the heroin-positive specimens were also positive for cocaine. For cocaine-positive specimens, in addition to parent cocaine, the confirmation process identified the following metabolic patterns: benzoylecgonine and cocaethylene (n = 31; 9.6 percent), cocaethylene only (n = 1; 0.3 percent), benzoylecgonine only (n = 59; 18.3 percent), and no metabolites (n = 20; 6.2 percent). Levels of cocaine detected in the confirmed samples ranged in value from 0.10 to over 100 ng/mg, with a median of 3.8 ng/mg. For heroin-positive specimens, both monoacyltl morphine and morphine were identified in the confirmation process in all but one of the specimens; in the remaining specimen, only morphine was identified. The median monoacyltl morphine metabolite level for the 12 specimens with this metabolite was 3 ng/mg, with a range of 0.33–6.1 ng/mg. We note that short segments (specimens less than 3 cm in length) constituted less than 10 percent of the specimens analyzed.

Sample prevalence rates by hair testing and interview report are listed in table 2. For cocaine, inpec-
tion of the McNemar chi-square test statistics indicates that hair testing yields a higher sample prevalence rate than do any of the survey-based measures of recent use. The sample prevalence estimate based on hair tests is more than 4.8 times the estimates generated from survey reports of past month use and 3.6 times the prevalence estimate generated from survey reports of past year use. The test statistics suggest that subjects were significantly less likely to disclose past month, past year, and lifetime use on the survey than they were to test positive for cocaine on the hair test.

Similar, albeit less marked, trends are observed for comparisons between the prevalence rate generated from heroin tests and rates derived from survey responses about heroin use. The test-based estimate is 2.2 times the rate generated from survey reports of past month day use and 1.6 times the rate generated from survey reports of past year use. McNemar tests indicate that differences between hair test results and survey reports of heroin use were not statistically significant. Of the three survey versus drug test heroin comparisons, only the comparison involving past month reports approached statistical significance.

Survey report and drug test overlap were explored in greater detail in table 3. Agreement between survey report and drug test results was particularly poor for cocaine, with none of the kappa or conditional kappa measures exceeding 0.20. Few respondents who reported recent (past month or past year) cocaine use had negative hair tests. Although survey report sensitivity clearly increased with more remote time periods for reporting, sensitivity of cocaine reporting was generally low, irrespective of survey measure. Recent cocaine (past month or past year) use was disclosed by fewer than one in five subjects who tested positive for this substance. Just over one in four subjects who tested positive for cocaine reported that they had ever used this substance.

There was a somewhat higher level of overlap between survey reports of heroin and hair test results. Kappa and conditional kappa coefficients for survey measures of heroin ranged from 0.29 to 0.43. Although sensitivities for heroin reports were somewhat higher than those for cocaine reports, overall rates of disclosure among those who tested positive were still low. Fewer than half of those who tested positive for heroin reported that they had ever used this substance.

Considering the high cocaine prevalence rate from hair tests, we evaluated the impact of raising the cutoff level for positive test results on our comparative analysis. We note that Kline et al. (9), following the suggestions of Baumgartner et al. (28), reported use of a cutoff corresponding to 0.2 ng/mg. Increasing the cutoff from 0.05 ng/mg to this value results in a net loss of five cocaine-positive cases and had minimal impact on the overall sensitivity of survey reports. Sensitivity was 19, 20, and 27 percent for past month, past year, and lifetime reports, respectively. A shift up to the 0.5 ng/mg cutoff used by Farabee and Fredlund (14) resulted in a net loss of 14 positive cases. Again, this upward shift had only a marginal impact on validity statistics, with sensitivity increasing to 21, 22, and 30 percent for past month, past year, and lifetime reports, respectively.

Toxicologists have suggested that the presence of metabolites at confirmation may be a better indicator of cocaine ingestion than the presence of parent cocaine alone (29). Accordingly, we reanalyzed comparisons after counting the 20 confirmed positive cases without metabolites as negatives. This change, as with the other changes noted above, had a marginal impact on sensitivity, with this statistic increasing to 21, 22, and 31 percent for past month and past year use and 31 percent for lifetime use. We did note, however, that kappa and conditional kappa values showed more than a mar-


<table>
<thead>
<tr>
<th>Survey measure</th>
<th>Specificity</th>
<th>SR-H-</th>
<th>Validity measure</th>
<th>SR-H+</th>
<th>K</th>
<th>Kt</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Sensitivity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cocaine</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Past month</td>
<td>98.6</td>
<td>208/211</td>
<td>18.0</td>
<td>20/111</td>
<td>0.20</td>
<td>0.12</td>
</tr>
<tr>
<td>Past year</td>
<td>96.3</td>
<td>201/211</td>
<td>18.9</td>
<td>21/111</td>
<td>0.17</td>
<td>0.10</td>
</tr>
<tr>
<td>Lifetime</td>
<td>81.5</td>
<td>172/211</td>
<td>27.0</td>
<td>30/111</td>
<td>0.09</td>
<td>0.07</td>
</tr>
<tr>
<td>Heroin</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Past month</td>
<td>99.4</td>
<td>307/309</td>
<td>30.8</td>
<td>4/13</td>
<td>0.41</td>
<td>0.30</td>
</tr>
<tr>
<td>Past year</td>
<td>98.4</td>
<td>304/309</td>
<td>30.8</td>
<td>4/13</td>
<td>0.34</td>
<td>0.29</td>
</tr>
<tr>
<td>Lifetime</td>
<td>96.1</td>
<td>297/309</td>
<td>46.2</td>
<td>6/13</td>
<td>0.36</td>
<td>0.43</td>
</tr>
</tbody>
</table>

* Survey report "no"/hair test negative.
† Survey report "yes"/hair test positive.
‡ Conditional kappa, with hair test as the standard.
original increase when the metabolite criteria were used. The kappas increased to 0.27, 0.22, and 0.14 (from 0.20, 0.17, and 0.09), and the conditional kappas increased to 0.16, 0.14, and 0.12 (from 0.12, 0.10, and 0.07) for past month, past year, and lifetime cocaine use comparisons, respectively. We noted, however, that these values were still well below the cutoff considered to be “fair” agreement (26).

Two potential limitations emerged from our survey report comparisons. Our recent use survey comparisons were limited to the periods past month and past year. The 3-cm segment that was typically analyzed corresponded to a period of 3 months of growth. This lack of overlap in use reference period potentially affected the overall levels of agreement between the two sources of use information. To investigate this issue, we examined the extent to which drug use reports for different use time frames were actually detected by the hair test (table 4). Although 87 percent of the self-reported past month cocaine users were detected by the hair test, only 12.5 percent of those whose reported use occurred between the past month and the past year were detected by the hair test. Similarly, 67 percent of self-reported heroin users for the past month were detected by the hair test, while none of those whose reported use occurred between the past month and past year were detected by the hair test.

In making these observations, we note that self-reported recent last use beyond the past month period was relatively rare for either cocaine or heroin. There were only eight reports of cocaine use and three reports of heroin use within this time frame. Most recent users reported that their last cocaine or heroin use occurred within the past month. We also note that while the potential lack of overlap in the detection period may affect the overall agreement statistics on table 3 (i.e., kappa and conditional kappa), it should not have affected the sensitivity rates. Use in the past 3 months is encompassed by past year use. Thus, if the drug test is accurate for the detection of use within a 3-month period, honest informants whose use was detected by the hair test should have disclosed use in a time frame encompassed by the past year.

Another potential limitation in our comparisons concerns the finding that nearly 10 percent of our specimens were below the length typically required for analysis. The analysis of shorter strands potentially limits the extent to which drug test insensitivity to self-reports of recent use was not prevalent in this study. Nevertheless, we evaluated the length of analyzed hair segments for those whose hair tests were negative but whose self-reports suggested recent cocaine and heroin use. Of the 10 self-reported recent cocaine users with negative hair tests, only two had a specimen with a length smaller than 3 cm. Of five self-reported recent heroin users with negative hair tests, none had a specimen with a length smaller than 3 cm. Thus, hair test insensitivity in this sample was probably not a consequence of insufficient hair length. In addition, we underscore that the major consequence of short specimens is on the overall levels of agreement (kappa and conditional kappa) and not on the sensitivity rate.

We investigated potential correlates of recent (past year) use disclosure among the subgroup of respondents who tested positive for cocaine (n = 111). Due to the small number of positive heroin tests, no similar comparisons were made for this substance. There were no associations between past year use disclosure and the age, education, employment status, race/ethnicity, or sex of the respondents. Disclosure was also not associated with interview condition. We do note, however, that respondent disclosure was associated with the amount of cocaine detected in the hair. Specifically, 34 percent of the respondents with levels of cocaine at or above the sample median value (3.8 ng/mg) disclosed past year use, compared with just 4 percent of respondents with levels below the median ($\chi^2_{df} = 1 = 16.6; p < 0.001$). Among 89 respondents testing positive for cocaine who did not report past year use, only five reported passive exposure. Of the 81 subjects who tested positive for cocaine but who

---

**TABLE 4. Hair test detection by survey reporting status by time frame, Chicago, Illinois, 1997**

| Drug use status (survey report) | Substance |
| --- | --- | --- | --- |
| | Cocaine | Heroin |
| | % positive | No. tested | % positive | No. tested |
| **Ever user** | | | | |
| Most recently reported use | | | | |
| Used in past month | 87.0 | 23 | 66.7 | 6 |
| Used in past year | 12.5 | 8 | 0.0 | 3 |
| Used >1 year ago | 23.7 | 38 | 22.2 | 9 |
| **Never user** | | | | |
| Total | 34.5 | 322 | 4.0 | 322 |

Am J Epidemiol Vol. 149, No. 10, 1999
reported never having used this substance, only four reported passive exposure.

Additional bivariate comparisons evaluating associations between demographic characteristics and lifetime drug use disclosure yielded two significant variables: race/ethnicity and education. Compared with those with more education, those with less than a high school education had higher rates of lifetime disclosure ($\chi^2_{df=2} = 8.64; p < 0.05$). Compared with Hispanic and black subjects, white/other subjects had higher rates of lifetime disclosure ($\chi^2_{df=2} = 6.07; p < 0.05$). In the context of multiple logistic regression simultaneously controlling for race/ethnicity, gender, education status, age, and study condition, only race/ethnicity maintained significance. Adjusted odds ratios suggested that, compared with black respondents who tested positive for cocaine, white/other respondents who tested positive had 1.9 times the odds of disclosing lifetime drug use (95 percent confidence interval 1.3–33.2); the contrast between Hispanic and white/other respondents was not significant.

**DISCUSSION**

The findings presented in this report suggest that survey reports of cocaine and heroin use do not provide an accurate picture of the prevalence of drug use in high-risk communities. Among those who consented to the hair test, drug use was underreported. Underreporting did not vary according to subject background characteristics or interview conditions. Disclosure did, however, vary by drug. It appears that heroin use was more likely to be disclosed than was cocaine use. In addition, those with higher concentrations of cocaine in their hair were more likely to disclose use of this substance. Finally, subjects who tested positive for cocaine were more willing to discuss lifetime drug use than recent use. These findings suggest that epidemiologic researchers should have more confidence in lifetime than in recent drug use prevalence rates. When recent drug use was reported, it was usually detected by the hair test. On the other hand, when recent drug use was detected by the test, it was usually not reported.

Findings with respect to the more accurate reporting of heroin use are consistent with at least two other studies of reporting behavior among drug users seeking treatment. Wish et al. (13) reported that clients who volunteered for drug treatment and who tested positive by hair for heroin nearly always reported recent heroin use at an initial intake screening; in contrast, clients who tested positive for cocaine reported recent use only about 79 percent of the time. While these authors suggest that treatment system incentives may be the cause of differences in disclosure rates, our findings derived in the context of a community sample suggest that reporting of heroin may be less sensitive than reporting of cocaine. Although the precise reasons for this need further exploration, we believe that these comparative findings support the observations of Magura and Kang (8) regarding the highly stigmatized nature of cocaine use, especially crack cocaine use, among young adults (30, 31).

Differential rates of disclosure with level of cocaine detected raise the possibility that our findings may potentially be an artifact of false positives among nonusers (i.e., passive exposure) among a significant portion of our sample. On the one hand, we note that few subjects in our sample who tested positive reported passive exposure. On the other hand, there is the possibility that subjects may be as reluctant to disclose their associations with drug users as they are to disclose their drug use; passive exposure may have been underreported here. Among the five cocaine-positive subjects who admitted to passive exposure but not to use, the median cocaine level was 0.91, a level well above the "high cutoff" used by Farabee and Fredlund (14). We also note that some toxicologists (20) point out that passive exposure cannot be differentiated according to levels of cocaine detected in the hair. We believe that the most accurate interpretation of our findings is that those who use cocaine heavily are more comfortable discussing their involvement with this substance than are more marginal users. This interpretation is consistent with findings from a previous analysis (23) indicating that respondents who reported lower levels of cocaine use in an initial interview were more likely to fail to report lifetime use at follow-up.

When heroin was detected by hair tests in our sample, cocaine was also almost always detected. This finding that heroin use in the community usually coincides with cocaine use is of considerable epidemiologic importance and provides empirical support to ethnographic accounts of the emerging phenomenon of "speedballing" (32). We have no way of knowing whether the actual use of these substances for the 11 subjects with co-occurring positive tests was simultaneous. Our study, like many epidemiologic surveys, considers heroin and cocaine as separate drug categories. These findings underscore the importance of incorporating survey measures specifically aimed at ascertaining whether heroin and cocaine were actually used in combination or separately.

The lack of association between demographic characteristics and recent drug use disclosure was somewhat surprising. The 111 subjects who tested positive for cocaine were demographically more homogeneous than was the rest of the sample, especially with respect to race/ethnicity. The subgroup that tested positive for
cigarette use was 80 percent black, 11 percent Hispanic, and only 9 percent white/other. Sample homogeneity, along with the reduced sample size, diminished the power to detect differences in disclosure rates across demographic subgroups. The finding suggesting that black informants were less likely to disclose lifetime use than were white/other informants, a finding that held up in a multivariate context, is consistent with recent literature discussing social differences with respect to drug disclosure (23–25). The contrast between recent and lifetime analyses reflects a general reluctance to disclose recent drug use among all subgroups testing positive for cocaine. Again, these findings support the notion that prevalence rates constructed from lifetime use reports have more validity than do rates derived from responses to questions about recent use.

Although the utility of hair testing has been demonstrated here and in other contexts, we underscore an important limitation of this procedure in community-based epidemiologic research. Just over half of the subjects in our study provided usable specimens. This participation rate is consistent with other work in samples not from corrections departments or treatment facilities (9). Nevertheless, the key demographic differences observed between participants and nonparticipants, especially with respect to race and gender, suggest that prevalence results based solely on this procedure may be subject to considerable bias.

Within the context of this study, there is the possibility that the $10 incentive motivated higher levels of drug test participation among drug users, especially nondisclosing drug users. Such differential selection into the test would yield comparative analyses that overestimate the extent of underreporting and drug use in the general population. Supplementary analyses suggested that there were no differences in the rates of disclosure of any recent and lifetime drug use between drug test participants and nonparticipants (33). Further, we note that even if every hair test nonparticipant was test negative for cocaine, the rate of cocaine use detected by the hair test for the entire sample would still greatly exceed rates of recent use derived from the survey. The hypothetical rate testing positive under this unlikely scenario would be 19.5 percent (111 of 568); this compares with 5.3 percent (30 of 568) reporting use in the past month, and 7.9 percent (45 of 568) reporting use in the past year. The accuracy of our estimates hinges both on the overall level of recruitment into the drug test and on respondent behavior (especially recent drug use) associated with test participation. Future studies need to more carefully assess the precise reasons and mechanisms associated with participation in hair and other drug testing procedures in community research.

ACKNOWLEDGMENTS

Supported in part by National Institute on Drug Abuse grants R01DA09285 and R01DA09286-02.

REFERENCES


Am J Epidemiol Vol. 149, No. 10, 1999