Introduction—The Validity of Self-Reported Drug Use: Improving the Accuracy of Survey Estimates

Lana Harrison and Arthur Hughes

ABSTRACT

Measuring levels and patterns of illicit drug use, their correlates, and related behaviors requires the use of self-report methods. However, the validity of self-reported data on sensitive and highly stigmatized behaviors such as drug use has been questioned. The goal of this monograph is to review current and cutting-edge research on the validity of self-reported drug use and to describe methodological advances designed to reduce total error in estimates of drug use and quantify sources of nonsampling error.

This monograph reviews a number of studies that use some presumably more accurate measure of drug use to validate self-reported use. In addition, evolving methods to improve a wide variety of procedures used in survey designs are explored, including computer-assisted interviewing, predictors of response propensity, measurement error models, and improved prevalence estimation techniques. Experimental manipulations of various survey conditions and situational factors also show promise in improving the validity of drug prevalence estimates in self-report surveys.

FOREWORD

The monograph arises from a technical review that was conducted on September 8 and 9, 1994 in Gaithersburg, MD, where papers were presented by 25 leading U.S. researchers on various aspects pertaining to the validity of self-reported drug use. The focus of the technical review was to examine recent research on validity using internal or external criteria, especially bioassays, as well as to examine methodological advances that can contribute to improved estimates of drug use in a survey environment. This monograph includes 20 of the 25 papers presented. The loss of several papers addressing the validity of the biological assays to assess drug use, particularly using

hair as the medium, are of particular concern. (Please refer to the Technical Note at the end of this Introduction.)

The Technical Review was broad based with two fairly distinct focuses. Hence, the first area of this monograph is an overview of what is known about the validity of self-report based on studies using internal and external validity criteria. Other chapters consider the importance of recanting earlier reports of drug use on longitudinal surveys and how ethnographic research methods may improve validity. The monograph includes overview chapters on several studies that attempt to determine the accuracy of self-reported drug use among criminal justice, treatment, and workplace populations by using urinalysis and/or hair analysis to validate recent drug use. Also included is a review article on the validity of biological assays to determine how accurately drug use is reported.

The second focus is on methodological advances that have been used or proposed as a means for understanding the extent of nonsampling error in surveys, and realizing further reductions in total error in estimates of drug use and associated behaviors. One promising method is the use of audio computer-assisted self-interviewing, which can allow complicated branching of questions to occur while permitting respondents with reading difficulties to complete the interview with minimum interviewer intervention. Other chapters deal with correlates of response propensity, cognitive laboratory procedures, privacy effects, sampling methods, measurement error models, and improved estimators of hardcore drug use. Overall, each chapter in this part of the monograph demonstrates where improvements can be achieved in the design of surveys collecting sensitive information, and how estimates of these behaviors can be improved.

As previously mentioned, the editors are concerned about the loss of several papers from the Technical Review that addressed the developing science of hair testing for drugs of abuse. These papers were based on the research of laboratory scientists and provided cautions about the state of the science with respect to the validity and reliability of methods to identify drugs in hair. The reader is referred to *Hair Testing for Drugs of Abuse: International Research on Standards and Technology*, edited by E. Cone, M. Welch, and M. Babecki, NIH Pub. No. 95-3727 (1995) for similar papers. The Technical Review and monograph include several papers detailing results of studies comparing drug use prevalence based on self-report, urinalysis, and hair testing measures. Since hair analysis is still a

developing science with unresolved issues, the results from these studies must be viewed with caution in light of the limitations of hair testing technology. (Please refer to the Technical Note at the end of this Introduction.)

Following is a review of the chapters in the order that they appear in the monograph.

MONOGRAPH OVERVIEW

Validity Studies

Harrison examines the research literature on validation studies to provide an overview of what is known about the accuracy of self-reported drug use. Before the mid-1980s, validation studies suggested that drug use was fairly accurately reported in self-report surveys. However, recent validation studies conducted with criminal justice and former treatment clients using improved urinalysis techniques and hair analyses suggest only about half or less of recent drug use is self-reported in confidential interviews. While this research has been used to criticize estimates of drug use generated from self-report surveys, there are limitations with the testing technology, as well as with the validity studies conducted to date. Harrison discusses these limitations, particularly with respect to urinalysis and the developing science of hair testing for drugs of abuse, but points out there is an accumulating body of research evidence that leads to some general conclusions about self-report. That is, self-report is less valid both for the more stigmatized drugs such as cocaine and for more recent rather than distant use. Self-report methods where respondents do not answer aloud increase reports of drug use. Also, the validity of self-report tends to be least reliable for those involved with the criminal justice system.

Former treatment clients, particularly narcotic users, have been the focus of much research on the validity of self-reported drug use. Harrell reports on a validity study conducted in 1985 in conjunction with the National Household Survey on Drug Abuse (NHSDA) sponsored by NIDA. NIDA chose not to publicly release the study because of an unacceptably low response rate. However, while the results must be viewed cautiously, they suggest variations in reporting by drug type, with the percentage of known users reporting their use highest for marijuana, followed by cocaine and hallucinogens, and lowest for heroin. The pattern of inconsistent reporting was

consistent with the social desirability hypothesis, with most admitting use of less stigmatized drugs but fewer admitting use of more stigmatized drugs.

In their chapter, Johnston and O'Malley examine the recanting of earlier reported lifetime use of several drugs from the Monitoring the Future study. Recanting rates are examined on nationally representative samples of high school seniors (18-year-olds) from the late 1970s as they are followed through age 32. Recanting rates were quite modest for the illegal drugs examined—marijuana, cocaine, and lysergic acid diethylamide (LSD)—but for the psychotherapeutic drugs examined (tranquilizers and barbiturates) they were more substantial. The larger differences for the psychotherapeutic drugs may be attributable to young adults correcting for earlier inaccurate reports of psychotherapeutic drug use due to difficulties in identifying the drugs. Consistent with earlier research, minorities—particularly African Americans—had somewhat higher rates of recanting on the illegal drugs. So did respondents in certain occupations, specifically the military and police/firefighting. In general, however, the evidence is quite good for validity of self-reported lifetime use of the illegal drugs gathered by mail in young adulthood.

The next chapter in the monograph presents an innovative approach to determining the reliability of self-reported drug use and drug dealing using both retrospective and prospective methods. Respondents were given a life history interview focusing on drug use history, involvement in drug sales, criminal history, violence history, and treatment history. They were also interviewed in detail about their activities over the past 7 days. Over the following 7 weeks, respondents were asked to report on activities in the preceding week. In general, Fendrich, Mackesy-Amiti, and Goldstein found the life history reports of current use for heroin, cocaine, marijuana, and alcohol were consistent with reports provided prospectively. However, subjects reported considerably higher use quantities and frequencies for substances in the life history reports than they did in the weekly interview reports. They also tended to underreport their alcohol use in the life history interviews compared to the weekly prospective interviews, suggesting they tended to minimize the importance of their alcohol use. However, with respect to heroin and cocaine, the phenomenon of overreporting was observed with respondents overestimating the volume and cost of cocaine and heroin they used in the life history interviews. On the other hand, nearly 20 percent of those reporting drug dealing in the weekly prospective interviews failed to report drug dealing in the life history

interviews. Preliminary inspection of the data suggests that some of the discrepancy in drug-dealing reports may be the result of discrepant definitions of dealing. This is especially applicable to low-level or sporadic dealers who, during the weekly prospective interviews, reported occasionally selling small quantities of drugs.

The next chapter, by Cone, assesses the strengths and limitations of biological assays to validate self-report. Over the past several decades, technologically sophisticated methods have been developed for analyzing drug metabolites in bodily fluids and tissues such as urine, blood, hair, saliva, semen, meconium, and perspiration. Each medium has advantages and disadvantages, and ongoing research is helping to further refine the tests. Drugs or their metabolites can generally be detected in urine for 2 to 4 days, although most illicit drugs are eliminated within 48 hours after use. Saliva offers advantages over urine, including a higher concen- tration of the parent drug than metabolites and a closer ratio to blood concentrations, but the window of detection is generally only 12 to 24 hours. Cone states this makes saliva most useful for the detection of recent drug use in accident victims, or testing employees before they engage in safety-sensitive activities. Research on sweat testing has been limited because of the difficulty of collecting sweat samples, but a sweat- collection device that is applied to the skin and worn for a period of several days to several weeks appears to have solved some of the collection problems and made sweat testing more feasible. The science of hair testing for drugs of abuse has improved in recent years, but there are still many unresolved issues. (Please refer to the Technical Note at the end of this Introduction.) Hair offers the potential for detecting drug use over much longer periods of time, which is very appealing; it can be easily stored and is less embarrassing to collect. Cone concludes that validation of self-report data by drug testing must be performed with careful consideration of the limitations imposed by the testing methodology and the biological specimen.

Preston, Silver, Schuster, and Cone discuss the innovative use of urinalysis to monitor treatment compliance in clinical trials. They report on their study of 37 patients who used cocaine consistently during the first 5 weeks of methadone treatment. Three days each week, subjects answered self-report questionnaires and submitted urine samples. Over the course of the 17-week clinical trial, subjects reported cocaine use on 20 percent of occasions, but tested positive for cocaine (qualitatively) on 68 percent of occasions. However, examination of the quantitative data reveals that at least part of the differential rates of self-report and qualitative cocaine-positive urine

specimen was due to carryover. A urine specimen collected several days after self-administration of a large amount of drug could have the same drug/metabolite concentration as a specimen collected just after self-administration of a small amount of drug. Concentrations of benzoylecgonine—a metabolite of cocaine—in urine specimens supported the suggestion that rates of drug use as determined by qualitative urinalysis were artificially high due to carryover. Preston and colleagues suggest that the effectiveness of substance abuse treatment programs can be monitored by frequently conducted urinalyses.

Miller, Donnelly, and Martz report on the forensic use of testing hair for drugs of abuse at the Federal Bureau of Investigation (FBI). Hair testing is only used by the FBI when other information exists that indicates drug use and the results can remove a person from suspicion or associate them with criminal activity. The detection of cocaine has been the FBI's first priority in hair testing for drugs of abuse because of its prevalence. Although the FBI does not routinely engage in testing hair for drugs of abuse, the chapter presents synopses of several cases where hair testing was used. Further, analysis of more than 100 samples was performed on hair obtained from a medical examiner's random autopsy collection. The results of the hair testing for drugs of abuse were found to be consistent with autopsy toxicology reports. Miller and colleagues conclude hair testing can be used in conjunction with urinalysis to give a more detailed drug history on a test subject.

In their chapter, Mieczkowski and Newel report on a study in which they compared urine and hair testing results among a population of Florida probationers. These probationers were already undergoing regular urinalysis, and were asked to participate in a confidential 6-month study that would also collect monthly hair samples. Of the 89 cases who had 6 complete sets of specimens, 36 were negative on all assays for all drugs, which was the most frequent finding. Focusing on the disconcordant hair and urine assay cases, the authors show that most of the disconcordant results were for cannabis and opiates. Mieczkowski and Newel have previously stated that hair testing is probably the best developed for cocaine, and their analysis helps to support their conclusion. They further posit that environmental contamination is not an unresolvable clinical problem for hair analysis of cocaine, provided one is willing to accept that marginal cocaine use, because of high cutoff values, may be classified as passive contamination. However, other research and researchers would disagree with their assertion this is resolvable with the current state of

technology (NIDA 1995). Mieczkowski and Newel caution that bioassays can create a false sense of certainty about the meaning and utility of biological testing of any kind. They suggest that hair assays should be used when the outcome cannot put the person undergoing the testing in jeopardy, and may be especially useful in epidemiological surveys.

Wish, Hoffman, and Nemes provide an overview of the research literature on the validity of self-report, making the point that as drug use became more stigmatized during the years of the War on Drugs, individuals may have become less willing to disclose past drug use. The research literature is replete with studies showing that individuals under criminal justice supervision are loath to report drug use on confidential and anonymous surveys. However, Wish and colleagues also suggest there is reason to question the validity of self-report among treatment clients—another group that has frequently been the focus of validity studies. Results are presented from a study of clients participating in the Washington, DC, Treatment Initiative study who were assessed for drug use by interview, urinalysis, and hair analysis. At intake, almost all clients who tested positive had reported their use of heroin (96 percent), but fewer clients had reported their cocaine use (82 percent). A subsample was followed posttreatment. Although information is not presented for urinalysis results, 62 percent tested hair-positive for opiates and 36 percent self-reported use, while 80 percent tested hair-positive for cocaine, and 52 percent self-reported their use in the past 90 days. One interesting finding was a strong association between the self-reported frequency of drug use and concentration of drugs found in the hair. Although this study can only be viewed as suggestive due to the limitations of hair testing technology and the small number of followup cases, Wish and colleagues assert that treatment evaluation studies that fail to validate their estimates of self-reported drug use should be interpreted with considerable caution. Clients may wish to show that the treatment they had participated in had some value.

In their chapter, Magura and Kang report the results of two validity studies conducted by the first author, one for a sample of patients in two methadone treatment programs in New York City and the other for a sample of criminally involved young adults. Self-report information and both urine and hair samples were obtained on all the clients. For the methadone sample, 60 percent self-reported recent cocaine use and

80 percent were hair positive. For the young adult sample, 23 percent self-reported recent cocaine use, but 67 percent were hair positive. Magura and Kang discount the sensitivity hypothesis because 75 percent reported lifetime drug dealing (41 percent in the past month). The curious finding, then, is the lower reports of cocaine use. Magura and Kang suggest that for the young adults, use of cocaine—or more specifically, crack—had become stigmatized, even though dealing of these drugs was not. However, there may be important explanations overlooked by the authors, based on the limitations of hair testing technology including issues of racial bias and passive contamination (see the Technical Note at the end of this Introduction).

Cook, Bernstein, and Andrews report on a study employing self-report, urinalysis, and hair analysis in a workplace sample. They selected a random sample of 1,200 employees of a steel plant in the western United States. Employees were randomly assigned to four different self-report methods of assessing illicit drug use: (1) individual interview in the workplace, (2) group-administered questionnaire in the workplace, (3) telephone interview, and (4) individual interview off the worksite. The group-administered questionnaire method produced prevalence rates that were roughly half those of the other self-report methods. However, perhaps surprisingly, Cook and colleagues found that self-reports produced higher prevalence rates than either urinalysis or hair analysis. For the entire sample, only 7.8 percent tested positive for any drug by urinalysis, while 9.4 percent reported recent drug use. For the subsample that had hair tests, 6.2 percent were positive for an illicit drug and 9.9 percent reported recent use. Nevertheless, Cook and colleagues found only about half of those positive for any drug on either test self-reported recent use. The authors concluded that the findings suggest the need for multiple assessment methods of estimating self-report. However, since most of those who tested positive by hair analysis were positive for marijuana, and hair analysis has been shown to be least reliable for detecting marijuana use, the need for multiple assessment methods does not appear a justifiable conclusion. In fact, the study results demonstrate that self-reports produced higher prevalence rates than either urinalysis or hair analysis.

Methodological Developments

Nonresponse error continues to be pervasive in surveys soliciting either sensitive or nonsensitive information. While surveys such as

NHSDA and Monitoring the Future typically achieve response rates from the upper 70s to mid 80s, little is known about what impact the nonrespon-dents (from 15 percent for high school seniors in Monitoring the Future to about 22 to 23 percent in NHSDA) have on estimates of drug use and other deviant behaviors. To gain a better understanding of nonresponse error in the NHSDA, Gfroerer, Lessler, and Parsley present results of the Census Match Study, a program where responding and nonresponding NHSDA households sampled in 1990 were matched to data from the 1990 Decennial Census. Information from the census on housing value, household composition, and other characteristics at the person, house-hold, block, and interviewer level were examined, with a subset of these variables found to be related to response propensity. This effort led to the development of improved nonresponse adjustment procedures in NHSDA. A second and unrelated study in this chapter called the Skip Pattern Experiment was fashioned to compare drug use reporting from two questionnaires: an experimental questionnaire that allowed the respondent to skip out of a set of questions if no drug use is reported, and the conventional questionnaire designed to require the respondent to answer all questions regardless of use. Results indicate that the skip pattern questionnaire produced less reporting of drug use.

Large-scale drug use surveys such as NHSDA provide excellent coverage of the general population and many demographic and socioeconomic subdomains; however, a sufficient number of sample members who use heroin regularly, for example, can be difficult to obtain using conventional sampling methods. Thompson's chapter presents innovative ways to reach sufficient numbers of these and other similar types of individuals through the use of adaptive sampling and graph sampling techniques. Also included is a discussion of resultant estimators that are design unbiased.

Understanding the methods used by researchers to measure the quality of self-reported drug abuse and associated behaviors is crucial. Hser's chapter provides a review of techniques used to assess reliability and validity of self-reported drug use and presents an assessment of the quality of self-report data among people at sexually transmitted disease clinics, emergency rooms, jails, and from a sample of narcotics addicts. Hser shows that adjustments for underreporting in these subpopulations should vary by gender, race, population type, and other factors. For example, among cocaine users who were self-reported nonusers, factors such as being female, minority, in jail, having multiple arrests in the past year, not being in treatment, and

being dependent in the past were significantly correlated with positive urine results.
Tourangeau, Jobe, Pratt, and Rasinski report findings from a methodo-logical study of reporting differences of sensitive behaviors such as pregnancy outcome (including abortion), the number of sexual partners, presence of a sexually transmitted disease, and level of condom use from a sample of women. Four modes of data collection by method of administration procedures were examined to determine the combination that results in higher levels of reporting. Overall, self-report clearly produced higher levels of reporting among women. Reporting based on use of computer-assisted collection versus conventional paper-and-pencil methods appear to be mixed.

In two studies, Lessler and O'Reilly compare the performance of audio computer-assisted self-interviewing (audio-CASI) with other methods such as an in-home computer-assisted personal interview (CAPI), out-of-home CAPI, and the traditional paper-and-pencil self-administered questionnaire (SAQ). In the first study, results show that computer-assisted interviewing produced higher rates of drug use reporting compared to the traditional SAQ procedure. The second study compared results of abortion reporting from the National Survey of Family Growth (NSFG), a major source of data on pregnancy and related information in the United States. Compared to CAPI, audio-CASI interviews produced reports of a higher number of abortions. Currently, more than 10,000 women participating in the NSFG have been successfully interviewed using audio-CASI technology.

When designing and conducting surveys involving sensitive topics such as drug use, it is important to have a good understanding of how privacy (or lack of it) during the interview affects the veracity of reporting. In a household survey of adults aged 18 to 45, Aquilino examined the effects of third-party presence on respondents' willingness to report drug use. Results show that the presence of a spouse or living partner while the interview was taking place did not seem to deteriorate the validity of self-report. On the other hand, truthful response appeared to decrease when a parent was present, even though all respondents were over age 17. In addition, these findings do not seem to vary by the three modes of administration used (self-administered, interviewer administered, and telephone).

Among other purposes, cognitive laboratory procedures can be used to gain a better understanding of how sensitive questions are perceived by the respondent. Willis provides a comprehensive and indepth review of the literature on cognitive laboratory-based research on sensitive

topics such as drug use, reproductive behavior, and drinking history. Based on laboratory research conducted by Willis and others, several recommendations are made related to the survey administration process. Some recommendations include the continued utilization of self-report as the primary mode of administration, shortening questionnaires on drug use, and limiting of complex concepts such as self-assessment of cause-and-effect relationships between drug use and deleterious life events.

Beimer and Witt provide a review of measurement error terminology such as measurement bias, reliability, validity, and mean square error. They present the mathematical relationship between reliability and validity (under appropriate assumptions), and discuss why measurement bias and validity should be treated as very different concepts. The main focus of this chapter is to examine the use of the Hui-Walter method for estimating measurement bias of self-reported drug use from the NHSDA. Taking advantage of redundancies in questions on drug use (i.e., lifetime use based on the recency question versus lifetime use based on any other question), the authors used this method to estimate false positive and false negative rates of drug use based on two sets of model assumptions: independence versus dependence of false negative rates between trials, among other things. A comparison of NHSDA false negative rates with denial rates from the National Longitudinal Survey of Youth (NLSY) generally showed, for example, a high correlation between the two for cocaine across various socioeconomic groups. This means that at a minimum, one will be able to determine which groups are more likely to contribute to false negative error.

Estimating hardcore use of drugs such as cocaine and heroin is a particularly challenging problem in major surveys due to the relatively small segment of the population involved in this behavior and the increased likelihood of underreporting very frequent use of these drugs. This will often lead to estimates with unacceptable sampling errors and measurement errors that may be much higher than those associated with other drugs and lower levels of use. In an attempt the address these shortcomings in the NHSDA, Wright, Gfroerer, and Epstein present a more sophisticated ratio estimator (than the one currently employed in the NHSDA) that incorporates population estimates from the Uniform Crime Report and the National Drug and Alcoholism Treatment Unit Survey. Results indicate that the alternative ratio estimator generated higher estimates of hardcore use of cocaine and heroin with higher levels of relative precision.

SUMMARY

In summary, the Technical Review sent a clear signal to the field that NIDA is supportive of the development and continuation of research on techniques to improve the validity of self-report and the accuracy of drug use estimates. Self-report will remain the primary mode of administration in drug use surveys, and is critical to obtain as valid information as possible. A developing body of important research information has been presented here and elsewhere about the successes and limitations of self-reported data collected from criminal justice and treatment populations. Researchers are also beginning to see validity studies conducted with general population groups, such as studies in the workplace. However, much more research needs to be conducted with more general popula-tions such as households and school students, the major source of drug use pattern and trend data for the Nation. There is also a growing body of research on validity studies that vary data-collection methodologies, but a much more systematic approach to determining the impact of various factors that may be manipulated in a survey environment needs to be employed.

With regard to improvements in sample design and estimation, it is hoped that at least some of these chapters will encourage those in the survey community to continue to pursue and develop better ways to collect sensitive data (e.g., via results obtained from laboratory procedures and through computerization), measure nonresponse error and measurement error in a quantitative manner, and develop better estimators. It is also hoped these chapters will encourage the design of improved survey procedures used on rare and hard-to-reach populations that result in significant reductions in both sampling and nonsampling error.

Appreciation goes out to the more than 100 individuals from the public and private sector who attended the 2-day technical review. Thanks also are due to Mary Beth Babecki, Marc Brodsky, James Colliver, Peter Delaney, Andrea Kopstein, and Elizabeth Lambert of NIDA, and Joseph Gfroerer and Doug Wright from the Substance Abuse and Mental Health Services Administration for serving as reviewers of earlier versions of selected chapters.

TECHNICAL NOTE: ISSUES PERTAINING TO THE USE OF HAIR AS A MEDIUM TO ANALYZE DRUG USE

The central focus of this monograph is the accuracy of self-reported drug use. Using hair as a medium to analyze drug use has been receiving increased attention because of the less embarrassing circumstances of collection, and because hair does not decompose like other body fluids/tissues. Hair testing also offers a wider time window of detection of drug exposure than conventional urine testing. However, a review of the current state of the science of hair testing technology demonstrates the unresolved issues with hair testing.

The first caution is that the studies using hair testing reported in this monograph are not state-of-the-art at the time of the publication of the monograph. Increased research on hair testing has led to even more questions about this developing science. The mechanism of how drugs enter the hair remains unknown (Cone and Wang 1995; Kidwell and Blank 1995). Understanding the pathway of drug entry into hair is important for interpretation of results, i.e., if drugs get into hair only from blood there is less risk of contamination and more likelihood of dose-concentration and time-location relationships existing; however, if sweat or sebum are important contributors, then these relationships are expected to be much less reliable and introduce the risk of environmental contamination. Research has demonstrated that passive contamination occurs, and that procedures to remove external contamination are not effective (cf., Kidwell and Blank 1995).

The basic pharmacological relationship between drug dose and concentration in hair has not been demonstrated; the amount of drugs incorporated into the hair depends on a variety of factors (Kidwell and Blank 1995). The relationship between time of drug exposure and location of drug in the hair strand has not been clearly established. Studies with labeled cocaine have found only a limited dose and time relationship (Cone 1994*a*; Henderson et al. 1993; Kidwell and Blank 1995).

There is considerable developing evidence that cocaine selectively accumulates in darkly colored (black) hair compared to brown or blonde hair (Cone 1994*b*; Henderson et al. 1993; Kidwell and Blank 1994). Commercial companies assert that their techniques “remove the melanin fraction” prior to analysis nullifying concerns about racial biases. However, NIDA research mimicking their techniques and in vitro binding experiments demonstrated a complete lack of

effectiveness for removing color bias from hair testing (Cone, personal communication, 1996).

Variability also exists across and within drug types in terms of the accuracy of detection. Cocaine has been shown to readily bind to hair (Kidwell and Blank 1994), but uptake and washout rates of cocaine in hair vary extensively between individuals. Hair testing for marijuana is the most difficult test to perform and the least reliable. Research shows much smaller amounts of cannabis are incorporated into the hair. Consequently, a positive finding for marijuana for a subject who denies use increases the likelihood that the report is not accurate, but should not be used as an absolute indicator. A negative test is even more unreliable and should not be used to conclude that marijuana was not used. Hair is not yet considered a good medium to test for cannabis use (Hindin et al. 1994).

Despite these limitations, hair is increasingly being used in prevalence studies as a measure of drug use. Several of these studies are reported in chapters in the monograph. These studies generally show higher rates of drug use obtained by hair analysis as compared to self-report. The chapters describing these studies may provide a few cautions about the science of hair testing, but results are generally presented as if the hair test results were totally accurate. The "absoluteness" of positive or negative findings is disputable. Since many of these studies included a large proportion of black subjects, the issue of racial bias in hair testing is especially salient to consider, as well as other limitations of the current state of hair testing technology. While the studies reported herein are clearly valuable and add to the accumulating knowledge on the developing science of hair testing, they must clearly be evaluated as suggestive rather than definitive. Although hair testing of subjects who provide self-report data increases the information base on these subjects, hair test results should not be regarded as the absolute reference criteria determining whether the subject is truthful or not.

Several controversial aspects of hair testing remain unresolved, although the technology has progressed rapidly over the last decade. Unfortunately, few clinical studies have been conducted that resolve important issues needed for interpretation of hair test results. The Office of Workplace Programs within the Substance Abuse and Mental Health Services Administration (SAMHSA) plans a Technical Review and resulting publication examining the current state of the science with respect to bioassay testing of bodily fluids and tissues in the spring of 1997. Hair testing will be a major focus. Readers should

refer to the resulting monograph for up-to-date information on hair testing technology.

REFERENCES

Cone, E. "Drug Testing in Hair." Paper presented at the SOFT Conference on Drug Testing in Hair, Tampa, FL, October 29-30, 1994*a*.

Cone, E. "Potential for Ethnic Bias in Hair Testing." Paper presented at the SOFT Conference on Drug Testing in Hair, Tampa, FL, October 29-30, 1994*b*.

Cone, E.J., and Wang, W.L. How environmental drug exposure can affect hair testing for drugs of abuse. In: Cone, E.; Welch, M.; and Babecki, M., eds. *Hair Testing for Drugs of Abuse: International Research on Standards and Technology*. USDHHS, NIDA, NIH Pub. No. 95-3727, 1995.

Henderson, G.L.; Harkey, M.R.; and Jones, R. Hair Analysis for Drugs of Abuse, Final Report, Grant No. NIJ90-NIJ-CX-0012. National Institute of Justice, National Institute on Drug Abuse, September 1993.

Hindin, R.; McCusker, J.; Vickers-Lahti, M.; Bigelow, C.; Garfield, F.; and Lewis, L. Radioimmunoassay of hair for determination of cocaine, heroin, and marijuana exposure: Comparison with self-report. *Int J Addict* 29(6):771-789, 1994.

Kidwell, D.A., and Blank, D.L. "Incorporation of Drugs of Abuse Into and Their Removal from Hair." Paper presented at the SOFT Conference on Drug Testing in Hair, Tampa, FL, October 29-30, 1994.

Kidwell, D.A., and Blank, D.L. Mechanisms of incorporation of drugs into a hair and the interpretation of hair analysis data. In: Cone, E.; Welch, M.; and Babecki, M., eds. *Hair Testing for Drugs of Abuse: International Research on Standards and Technology*. USDHHS, NIDA, NIH Pub. No. 95-3727, 1995.

National Institute on Drug Abuse. Cone, E.; Welch, M.; and Babecki, M., eds. *Hair Testing for Drugs of Abuse: International Research on Standards and Technology.* USDHHS, NIDA, NIH Pub. No. 95- 3727, 1995.

AUTHORS

Lana Harrison, Ph.D.
Associate Director
Center for Drug and Alcohol Studies
University of Delaware
77 East Main Street
Newark, DE 19716

(Formerly Statistician
Division of Epidemiology and Prevention Research
National Institute on Drug Abuse)

Arthur Hughes, M.S.
Chief
Epidemiology Research Branch
Division of Epidemiology and Prevention Research
National Institute on Drug Abuse
Room 9A-53
5600 Fishers Lane
Rockville, MD 20857

Click here to go to page 17