National Survey on Drug Use and Health: Summary of Methodological Studies, 1971–2005

DEPARTMENT OF HEALTH AND HUMAN SERVICES
Substance Abuse and Mental Health Services Administration
Office of Applied Studies

Acknowledgments

This report was prepared by the Office of Applied Studies (OAS), Substance Abuse and Mental Health Services Administration (SAMHSA), and by RTI International, a trade name of Research Triangle Institute, Research Triangle Park, North Carolina. Work by RTI was performed under Contract No. 283-03-9028. Contributors and reviewers at RTI listed alphabetically include Tim Flanigan, Mary Ellen Marsden, Emily McFarlane, Joe Murphy, Hyunjoo Park, Lanny L. Piper, Adam Safir, Thomas G. Virag (Project Director), and Kevin Wang. Contributors at SAMHSA listed alphabetically include Joseph Gfroerer and Arthur Hughes. At RTI, Richard S. Straw edited the report with assistance from Jason Guder and Claudia M. Clark. Also at RTI, Roxanne Snaauw formatted and word processed the report, and Pamela Couch Prevatt, Teresa F. Gurley, and Shari B. Lambert prepared its press and Web versions. Final report production was provided by Beatrice Rouse, Coleen Sanderson, and Jane Feldmann at SAMHSA.

Obtaining Additional Copies of Publication

Copies may be obtained, free of charge, from the National Clearinghouse for Alcohol and Drug Information (NCADI), a service of SAMHSA. Write or call NCADI at:

National Clearinghouse for Alcohol and Drug Information
P.O. Box 2345, Rockville, MD 20847-2345
1-240-221-4017, 1-800-729-6686, TDD 1-800-487-4889

Electronic Access to Publication

This publication can be accessed electronically through the Internet connections listed below:

http://www.samhsa.gov
http://www.oas.samhsa.gov

Originating Office

SAMHSA, Office of Applied Studies
1 Choke Cherry Road, Room 7-1044
Rockville, MD 20857

September 2006

Table of Contents

Introduction

Purpose of Report

This report provides a comprehensive summary of methodological research conducted on the National Survey on Drug Use and Health (NSDUH). Since its inception in 1971, considerable research has been conducted on the methods used in the survey. Studies have addressed sampling, counting and listing operations, questionnaire design, data collection procedures, editing and imputation procedures, nonresponse, and statistical analysis techniques associated with the survey. This research has been critical to the NSDUH project, providing information to identify and quantify survey errors, to guide redesign efforts, and to develop more efficient survey processes. These research studies have been published in professional journals, books and book chapters, conference proceedings of professional association meetings, and reports published by Federal agencies, such as the Substance Abuse and Mental Health Services Administration (SAMHSA) and the National Institute on Drug Abuse (NIDA).

The summaries of published studies included in this report provide a record of the wealth of methodological findings stemming from NSDUH research. This research is of great value to users of NSDUH data who need to understand the impact of method effects on their own studies. Researchers using NSDUH data for epidemiological studies need to understand these effects when designing their analysis plans and also should be aware of these issues in interpreting their results. The NSDUH body of methodological research also will be useful to those responsible for managing large-scale surveys or designing new surveys, particularly if they will collect substance use data. Compiling all of these NSDUH-related findings in one document will make it much easier for these researchers to find the studies most relevant for their purposes.

Structure

The report is organized into chapters based on year of publication. Within each chapter, publication summaries are sorted alphabetically by the last name of the first author. For increased readability, publication summaries are further organized into four uniform sections. The Citation section provides bibliographic detail. The Purpose/Overview section presents the motivation for the research and reviews the history or background of the subject matter, if available. For experiments or other analyses, the Methods section describes the methods used to conduct the research; for reports covering multiple method initiatives, or that describe or summarize the results of other studies, the Methods section presents an overview of the organization of the publication. The Results/Conclusions section summarizes the findings. When the nature of the publication is such that there are no methods or results/conclusions per se (e.g., introductory comments to an invited paper session published in a proceedings volume), an "N/A" appears in the section.

For publications in the public domain, such as those prepared by RTI International[1] (the NSDUH contractor since 1988) or SAMHSA, the original verbatim abstract was used as a starting point for the final edited summary. Summaries of copyrighted publications, such as

[1] RTI International is a trade name of Research Triangle Institute.

books or journal articles, were written independently of existing abstracts after substantial review of the source material. In all cases, the summaries in this report present mainly key points from source publications. Therefore, readers are encouraged to access the source publication for complete information.

Sources

The publications summarized in this report were compiled through an extensive search of archives, databases, books, and reports. A list of the sources used to prepare the report is provided below (note: some sources do not go back as far as 1971). The source publications for much of the research summarized in this report are available online and can be found by visiting the websites listed in the Archives/Databases section below.

Archives/Databases

- Conference Programs of the American Association for Public Opinion Research (AAPOR)
  http://www.aapor.org/default.asp?page=conference_and_events/past_conferences
- Proceedings of the Survey Research Methods Section of the American Statistical Association (ASA)
  http://www.amstat.org/sections/srms/Proceedings/
- Catalog of U.S. Government Publications (CGP)
  http://catalog.gpo.gov/F
- Proceedings of the Conference on Health Survey Research Methods (CHSRM)
  http://www.cdc.gov/nchs/
- Entrez, the Life Sciences Search Engine (National Center for Biotechnology Information)
  http://www.ncbi.nlm.nih.gov/gquery/gquery.fcgi
- Education Resource Information Center (ERIC)
  http://www.eric.ed.gov/ERICWebPortal/Home.portal
- Government Accountability Office (GAO)
  http://www.gao.gov/
- Scholarly Journal Archive (JSTOR)
  http://www.jstor.org/
- MathSciNet
  http://www.ams.org/mathscinet
- *Medical Care*, Journal of the Medical Care Section, American Public Health Association
  http://www.lww-medicalcare.com/pt/re/medcare/home.htm

- National Academies Press
  http://www.nap.edu/

- National Clearinghouse for Alcohol and Drug Information (NCADI)
  http://ncadi.samhsa.gov/research/default.aspx?

- National Criminal Justice Reference Service (NCJRS) Abstracts Database
  http://www.ncjrs.gov/abstractdb/search.asp

- National Institute on Drug Abuse (NIDA)
  http://www.nida.nih.gov/

- National Survey on Drug Use and Health (NSDUH)
  http://www.oas.samhsa.gov/nsduh.htm

- PsycINFO
  http://www.apa.org/psycinfo/

- PubMed (National Library of Medicine)
  http://www.ncbi.nlm.nih.gov/entrez/query.fcgi

- Substance Abuse and Mental Health Data Archive (SAMHDA)
  http://www.icpsr.umich.edu/SAMHDA/

- Sociological Abstracts
  http://www.csa.com/factsheets/socioabs-set-c.php

- Web of Science (includes Arts and Humanities Citation Index, Science Citation Index Expanded, and Social Sciences Citation Index)
  http://www.scientific.thomson.com/products/wos/

- WorldCat (combined book catalog of over 30,000 libraries)
  http://www.oclc.org/worldcat/default.htm

Publications Archived at SAMHSA

- 1992 National Household Survey on Drug Abuse: Findings of first quarter skip test: Final report
  http://www.oas.samhsa.gov/methodsHY/SkipTest92.pdf

- The effectiveness of cash incentives to respondents as an inducement to cooperate in completing an in-home interview: Fall 1981 incentive evaluation report for the National Drug Abuse Survey
  http://www.oas.samhsa.gov/methodsHY/CashIncentives81.pdf

- Estimates of drug use prevalence in Miami from the 1991-1994 National Household Survey on Drug Abuse: Methodological report
  http://www.oas.samhsa.gov/methodsHY/Miami9193.pdf

- Evaluation of results from the 1992 National Household Survey on Drug Abuse (NHSDA)
  http://www.oas.samhsa.gov/methodsHY/PeerReview92.pdf

- Report on 1990 NHSDA-Census Match
  http://www.oas.samhsa.gov/methodsHY/Census90.pdf

- Sensitivity of survey techniques in measuring illicit drug use
  http://www.oas.samhsa.gov/methodsHY/Validity75.pdf

- The validity of self-reported drug use data: The accuracy of responses on confidential self-administered answer sheets
  http://www.oas.samhsa.gov/methodsHY/Validity86.pdf

Books

- Survey Measurement of Drug Use: Methodological Studies, edited by Turner, Lessler, and Gfroerer

- Redesigning an Ongoing National Household Survey: Methodological Issues, edited by Gfroerer, Eyerman, and Chromy

- Evaluating and Improving Methods Used in NSDUH, edited by Kennet and Gfroerer

Reports

- Results from the 2003 NSDUH: National Findings

- Results from the 2002 NSDUH: National Findings

In addition to the sources mentioned above, internal RTI lists of conference papers, dissertation abstracts, and staff resumes were reviewed for NSDUH methodological citations. Internet search engines, such as Google.com, iTools.com, and Scirus.com, also were used to find citations of interest.

Survey Background and History of Methodological Studies

NSDUH is an annual survey of the civilian, noninstitutionalized population of the United States aged 12 years old or older. Prior to 2002, the survey was called the National Household Survey on Drug Abuse (NHSDA). NSDUH is the primary source of statistical information on the use of illegal drugs by the U.S. population. Conducted by the Federal Government since 1971, the survey collects data by administering questionnaires to a representative sample of the population through face-to-face interviews at their places of residence. The survey is sponsored by SAMHSA of the U.S. Department of Health and Human Services and is planned and managed by SAMHSA's Office of Applied Studies (OAS). Data collection is currently conducted under contract with RTI International, Research Triangle Park, North Carolina.

The first two surveys, done in 1971 and 1972, were done under the auspices of the National Commission on Marihuana and Drug Abuse. These initial surveys of about 3,000 respondents each year employed data collection procedures designed to elicit truthful responses to questions on drug use. The methodology included the use of anonymous, self-administered answer sheets completed in a private place within respondents' homes. The NHSDA was continued on a periodic basis by NIDA, starting in 1974, using the same basic design. As the demand for data and the importance of the survey increased, the sample size was increased, and the questionnaire and methodology were refined. In particular, the increasing impact of cocaine abuse in the United States in the 1980s led to significant increases in the sample, driven in part by the needs of the newly created Office of National Drug Control Policy, which required improved data for tracking illicit drug use. With the larger survey cost and more prominent role of NHSDA data among policymakers, the need for more research on the NHSDA methods was clear. Thus, beginning in the late 1980s, methodological research became an ongoing component of the NHSDA project under NIDA. When the project was transferred to OAS in the newly created SAMHSA in 1992, the methodology program continued.

The initial series of studies (most were published in Turner, Lessler, & Gfroerer, 1992) evaluated the basic survey methods and tested possible improvements, ultimately leading to a redesign of the questionnaire in 1994, using the same basic data collection method (self-administered paper answer sheets). Soon after this redesign, however, SAMHSA decided in 1995 to pursue the use of a newly emerging data collection technology, audio computer-assisted self-interviewing (ACASI). Thus, NHSDA methodological research turned to the testing and development of ACASI, which SAMHSA ultimately implemented in 1999 (OAS, 2001). The ACASI implementation occurred simultaneously with a new sampling plan, designed to produce State-level estimates, with a sample size of about 70,000 respondents per year. The implementation of the 1999 redesign generated a number of methodological results, stemming both from the development of the new design and from the assessment of its impact (Gfroerer, Eyerman, & Chromy, 2002). Other interesting methodological findings emerged following the 1999 redesign; in 2002 and beyond, with the testing and implementation of respondent incentives and when the name of the survey was changed to the National Survey on Drug Use and Health; and with other assessments of the survey procedures (Kennet & Gfroerer, 2005).

This report will be updated on a regular basis to include summaries for new method publications. Although every effort has been made to include citations for all method studies published to date, it is possible that some have been omitted. Readers are encouraged to notify OAS about any omissions by visiting http://www.oas.samhsa.gov/Mail/email.cfm.

For inquiries on the availability of particular NSDUH research and other questions, please visit the OAS webpage at http://www.oas.samhsa.gov/nhsda.htm. To order single volumes in the NSDUH Series or the Methodology Series of reports or to be put on the mailing list to receive copies of reports when they become available, visit:

- http://www.oas.samhsa.gov/mail/publst.cfm
- http://www.oas.samhsa.gov/mail/maillst.cfm

1975–1986

The effectiveness of cash incentives to respondents as an inducement to cooperate in completing an in-home interview: Fall 1981 incentive evaluation report for the National Drug Abuse Survey

CITATION: Abelson, H. I., & Fishburne, P. M. (1981, December). *The effectiveness of cash incentives to respondents as an inducement to cooperate in completing an in-home interview: Fall 1981 incentive evaluation report for the National Drug Abuse Survey* (Report No. RAC 4305 prepared for the Social Research Group at George Washington University and the National Institute on Drug Abuse). Rockville, MD: National Institute on Drug Abuse.

PURPOSE/OVERVIEW: In the late 1970s, survey organizations reported increasing difficulty in obtaining respondent cooperation. A growing body of literature suggested that one means of reducing respondent noncooperation was to use an incentive (i.e., a reimbursement to respondents for the burden of responding). Such incentives were seen as a means of enhancing respondent motivation, reducing respondents' perceptions of inconvenience, reducing item nonresponse, and improving interviewers' expectations of respondent cooperation. In surveys prior to the 1982 National Household Survey on Drug Abuse (NHSDA), respondents received a gratuity valued at approximately $3.00. This practice was continued in the 1982 survey. In an effort to determine the cost-effectiveness of this practice in reducing interviewer callbacks and ensuring at least an 80 percent response rate, an incentive evaluation study was undertaken as part of the 1982 survey.

METHODS: During the initial 30-day field data collection period, the contractor conducted an evaluation of the incentive payment to the respondents. A cluster of approximately 300 households was assigned to different conditions. The control conditions were defined as no payment to respondents. Two experimental conditions (approximately 100 respondents each) were defined in terms of size of payment. Each of the experimental groups received $3.00 and $6.00, respectively. The authors performed an empirical test of the effectiveness and cost-efficiency of incentives. Specifically, the authors compared the use of $3.00 and $6.00 incentives with the use of no incentive to test the hypotheses that (1) the use of incentives increases the response rate, (2) the use of incentives decreases the average number of visits to the households, and (3) the use of incentives decreases the average direct cost per completed interview.

RESULTS/CONCLUSIONS: The data showed no statistically significant differences attributable to the incentive treatments. Specifically, with regard to the three hypotheses tested, there were no statistically significant differences between the three treatment levels in terms of response rate, average number of visits to the household, or average direct cost per completed interview.

Sensitivity of survey techniques in measuring illicit drug use

CITATION: Cisin, I. H., & Parry, H. J. (1975, March). *Sensitivity of survey techniques in measuring illicit drug use* (Report No. 75-1, Contract No. HSM 42-73-197 (ND)). Rockville, MD: National Institute on Drug Abuse.

PURPOSE/OVERVIEW: By the mid-1970s, the problem of validity of response was one that had troubled sample survey researchers ever since the great expansion of the field in the 1930s. A considerable number of methodological studies, many of the most ingenious kind, had addressed themselves to the query: "How much credence can we give to behavior reported by respondents, particularly when self-esteem or deviant behavior is involved?" Most studies of validity emphasized the expectation that social pressure on respondents may result in the underreporting of behavior that could be called deviant or at least inappropriate and the overreporting of behavior that can add to self-esteem. Sometimes, of course, researchers for one reason or another may miscalculate the amount and direction of social pressure to which their respondents are subjected. Kinsey, for example, in his pioneering study of male sexuality was laudably desirous of steering at safe distance from the Scylla of underreporting certain types of sexual behavior. In his attempt to avoid an obvious danger, in the view of some critics, he passed too close to the Charybdis of overreporting.

Methodological studies in the field of validity of response had, in general, been of two kinds by the mid-1970s. The first type consisted of a comparison of two groups: an index group that because of known or assumed past behavior was significantly more likely than a control group to have acted in a certain way. A survey conducted among matched samples of the two groups should yield results showing significantly greater amounts of the behavior being studied among the index group. The second type of validity study was where reported behavior of a "known" group is compared with actual previously recorded behavior. Prior to the full-scale national survey on drug abuse being conducted in 1975, it was decided to precede the survey with a special validity study to examine the responses of a "known" index group with those of a control group. Additionally, after the validity study went into the field, the authors decided to augment it with a comparison of survey responses with records.

METHODS: The index group consisted of 100 respondents who had sought and received treatments at certain drug abuse clinics in the San Francisco Bay area. The control group consisted of 100 randomly selected respondents who were matched against a member of the index group in terms of age, gender, and location, but who had not sought help at a drug abuse clinic. Each group was asked a series of questions concerning a large number of drugs, questions dealing with past use and current use, with the influence of the peer group, and with past year incidence (i.e., use of a drug for the first time in the 12 months preceding the survey). Responses of the two groups then were compared, both in terms of individual drug behavior and in use of various combinations of drugs.

RESULTS/CONCLUSIONS: The study found a large number of significant differences in drug behavior between the two groups. Such differences tended to be concentrated at the beginning of the drug use continuum (i.e., the "used at some time" column). Much less difference was found in the "current use" percentage partly because current use for any given drug (except marijuana)

was generally low, so that differences were not statistically significant. Nevertheless, in nearly all cases the differences were in the expected direction.

Heroin incidence: A trend comparison between national household survey data and indicator data

CITATION: Crider, R. A. (1985). Heroin incidence: A trend comparison between national household survey data and indicator data. In B. A. Rouse, N. J. Kozel, & L. G. Richards (Eds.), *Self-report methods of estimating drug use: Meeting current challenges to validity* (pp. 125-140). Rockville, MD: National Institute on Drug Abuse.

PURPOSE/OVERVIEW: A commonly held belief in the mid-1980s was that respondents in a face-to-face survey will underreport the extent of their involvement in deviant behavior (Fishburn, 1980; Miller, 1985 [see abstract in this volume]; Sirkin, 1975). To test this assumption, a time series of self-reported year of first use of heroin was compared with a time series of heroin indicator data, such as heroin-related emergencies and deaths, treatment admissions, and hepatitis B cases. Although there was a general belief that self-reported heroin use would be underreported, no test of that assumption had been conducted comparing the time-series trend of heroin incidence based on survey data to the time-series trends of heroin indicators.

METHODS: Seven drug indicators were compared: (1) the number of heroin initiates based on the face-to-face National Household Survey on Drug Abuse (National Survey); (2) the number of heroin initiates voluntarily entering a panel of consistently reporting federally funded treatment programs (Client-Oriented Data Acquisition Process, CODAP) for the first time by the year of first use; (3) the residual number of hepatitis B cases per year; (4) the percentage of high school seniors ever using heroin (Johnston et al., 1982); (5) the number of heroin-related emergency room visits reported to the Drug Abuse Warning Network (DAWN); (6) the number of heroin-related deaths reported to DAWN; and (7) the average street-level heroin purity (U.S. Drug Enforcement Administration [DEA], 1984). Because of the small proportion of the population reporting ever having used heroin, the year of first use data from the National Institute on Drug Abuse's (NIDA's) National Surveys conducted in 1977, 1979, and 1982 were pooled to show the number of new users in the household population by year of first use. In addition, the data were "smoothed" by using a 2-year moving average.

RESULTS/CONCLUSIONS: The number of cases of heroin use combined from the three National Surveys were shown by year of first use, 1965 through 1980. Although the number of heroin users in any one survey was small, the pooled data produced frequencies large enough to establish trends. These data were not used to make estimates of the number of heroin initiates in any particular year, but were used to show a changing pattern over a several-year period. The epidemics in the early 1970s and the mid-1970s were evident. These epidemic periods occurred at the same periods reported by high school seniors and by heroin users in treatment. The household self-report data trends based on age and frequency of use also were consistent with the trends in periods of initiation reported by heroin users in treatment as noted in drug abuse treatment admission data for year of first heroin use. Trends in indicators of heroin epidemics were compared with trends based on self-report data from the National Surveys. The trends in

hepatitis B cases, heroin-related emergency room visits, heroin-related deaths, and the average retail heroin purity were consistent with the epidemic periods suggested by the household data. This consistency among the three sources of self-reported data on trends in year of first heroin use, combined with the consistency of these self-reported data with the trends based on the indicators of heroin epidemics, offers some validation to the use of retrospective direct questions concerning age of first use of heroin to monitor heroin incidence in the household population.

Influence of privacy on self-reported drug use by youths

CITATION: Gfroerer, J. (1985). Influence of privacy on self-reported drug use by youths. In B. A. Rouse, N. J. Kozel, & L. G. Richards (Eds.), *Self-report methods of estimating drug use: Meeting current challenges to validity* (DHHS Publication ADM 85-1402, NIDA Research Monograph 57, pp. 22-30). Rockville, MD: National Institute on Drug Abuse.

PURPOSE/OVERVIEW: Underreporting of drug use by survey respondents has always been a major concern of drug abuse survey researchers. This concern particularly applies when the respondents are youths because they might fear being punished if their use of drugs was discovered by parents. By 1985, a number of studies had been conducted to determine whether underreporting was a serious problem and to identify procedures that could be used to obtain the most valid data. Although some of the results were contradictory, most studies concluded that reliable, valid self-reported drug use data could be obtained (Hubbard et al., 1976; O'Malley et al., 1983; Single et al., 1975; Smart, 1975; Smart & Jarvis, 1981). Factors identified as possibly affecting the reporting of drug use by youths included the type of questionnaire (interview vs. self-administered), characteristics of the interviewer, the degree of anonymity of the respondent, the setting (home vs. school), and the degree of privacy during the interview (Johnston, in press; Sudman & Bradburn, 1974). In several studies, self-administered questionnaires were shown to produce higher reported prevalence of drug use than interviews. Although many factors that may affect underreporting could be controlled by researchers, it was not possible to achieve complete privacy (i.e., no other person in the room) in every interview when conducting a household survey. Given this limitation, it was important to assess the impact of the lack of privacy on the results of a survey, both to assess the potential impact on the validity of data from that survey and also to provide general information on the importance of privacy for future surveys.

METHODS: This study used data from a national probability sample of youths 12 to 17 years of age to assess the impact of the lack of privacy on the results of surveys. Depending on household composition, interviews were conducted with one adult only, one youth only, or both an adult and a youth. The surveys collected data on whether respondents had used various licit and illicit drugs in the past month, past year, or ever. Measures were taken to ensure privacy and confidentiality.

RESULTS/CONCLUSIONS: It was found that interviewers were successful in obtaining privacy, and the degree of privacy obtained was the same in a 1979 interview as in a 1982 interview. In general, it was found that most population groups that reported higher prevalence of drug use did not have significantly more privacy than lower prevalence groups. One exception was whites, who had slightly higher reported prevalence of cigarette and alcohol use than other races and also had more privacy during interviews. An investigation of this indicated that the

privacy effect is independent of race. In conclusion, it appeared that reporting of drug use by youths was affected by the degree of privacy during the interview, even when a self-administered answer sheet was used by the respondent. This underscored the importance of achieving maximum privacy when conducting drug surveys and raised questions regarding the validity of data from surveys in which adequate privacy was not obtained.

The validity of self-reported drug use data: The accuracy of responses on confidential self-administered answer sheets

CITATION: Harrell, A. V., Kapsak, K. A., Cisin, I. H., & Wirtz, P. W. (1986, December). *The validity of self-reported drug use data: The accuracy of responses on confidential self-administered answer sheets* (Contract No. 271-85-8305). Rockville, MD: National Institute on Drug Abuse.

PURPOSE/OVERVIEW: The purpose of this study was to examine the validity of self-reported data on the use of illicit drugs, including marijuana, cocaine, hallucinogens, and heroin, as well as the nonmedical use of some psychotherapeutic drugs. The focus of the study was the criterion validity of self-reported data and the extent to which respondent accounts of their drug use could be shown to be consistent with an independent indicator (i.e., clinic records of respondents' drug use).

METHODS: To examine the validity of self-reported drug use data, a sample of clients discharged from drug treatment facilities was interviewed using a questionnaire and interviewing procedures developed for the 1985 National Household Survey on Drug Abuse (NHSDA). Respondents' self-reported drug use was compared with the criterion of clinic records on the drugs being abused by these same respondents at the time of their admission to treatment. To isolate relatively recent drug use that would be less subject to failures of recall, the sample included only respondents who had been admitted to treatment within a year of the interview. To avoid possible bias in the drug use reports that could occur if respondents were aware that their answers could be verified, respondents were treated exactly as though they had been randomly selected to participate in NHSDA, and no reference was made to any advance knowledge of their drug use history. Similarly, interviewers were unaware of the fact that respondents had been treated for drug abuse, both to protect the privacy of the respondents and to prevent interviewer expectations from influencing responses to the questionnaire.

The interview provided self-reported data on the use of a wide variety of illicit drugs and on respondents' symptoms of, and treatment for, drug abuse. The clinic records provided data on the drugs (up to three) that were considered problems at the time of admission to treatment. The validity "tests" were whether respondents reported on the survey use of the drugs listed in the clinic records at the time of admission in the past year or at any time in the past. Another validity "test" was whether the respondents reported receiving any treatment for the use of drugs other than alcohol during the past year or at any time in the past. The analysis investigated the relationship of reporting accuracy to demographic characteristics and to factors that might influence the respondents' willingness or ability to respond accurately, such as the level of privacy during the interview and the amount of time between admission to the program and the interview.

RESULTS/CONCLUSIONS: In general, the reporting pattern on the drug use questions was consistent with the thesis that the higher the level of social stigma associated with a behavior, the more prevalent denial of the behavior will be. Variation in the level of privacy during the interview did not appear to influence the willingness of respondents to report drug use. Responses to questions about drug treatment were generally less accurate than the drug use self-reported data.

The nominative technique: A new method of estimating heroin prevalence

CITATION: Miller, J. D. (1985). The nominative technique: A new method of estimating heroin prevalence. In B. A. Rouse, N. J. Kozel, & L. G. Richards (Eds.), *Self-report methods of estimating drug use: Meeting current challenges to validit*y (DHHS Publication No. ADM 85-1402, NIDA Research Monograph 57, pp. 104-124). Rockville, MD: National Institute on Drug Abuse.

PURPOSE/OVERVIEW: In 1985, the nominative technique was a relatively new method of indirect survey-based estimation that was being developed expressly for the purpose of estimating heroin prevalence in the general population. This technique, which involves asking respondents to report on their close friend's heroin use, is essentially an attempt to reap the benefits of survey research, while at the same time avoiding some of the major problems of the self-report method. The primary purpose of the nominative technique is to minimize respondent denial of socially undesirable behavior. Another possible advantage is achieving coverage of "hard-to-reach" deviant population groups. The nominative question series was inserted in the 1977, 1979, and 1982 National Household Surveys on Drug Abuse (NHSDAs) (Miller et al., 1983).

METHODS: The resulting nominative estimates of heroin prevalence are presented and contrasted with corresponding self-report estimates.

RESULTS/CONCLUSIONS: Nominative estimates of heroin prevalence have been consistently higher than self-reports of heroin use. During this time, nominative data have generally followed mainstream patterns of drug use: Nominative estimates for young adults and for males were higher than nominative estimates for older persons, youths, and females. Moreover, in 1985, the recent downward trends in drug use were replicated by the nominative heroin data. Thus, the overall picture presented by the nominative data—similar patterns but higher levels of prevalence—seemed to support the validity of the new approach. Nevertheless, considerable caution should be exercised in interpreting nominative data. This is chiefly because a substantial minority of nominators cannot report the number of other close friends of the heroin user who also "know." Although missing data have been handled by a conservative imputation rule, the fact that so many persons are unable to provide an answer to this key question casts doubt on the accuracy of the answers that were given. In fact, the nominative approach might tend to produce overestimates because of the potential for undercounts of the numbers of others who "know."

1990–1991

Generalized standard error models for proportions in complex design surveys

CITATION: Bieler, G. S., & Williams, R. L. (1990). Generalized standard error models for proportions in complex design surveys. In *Proceedings of the 1990 American Statistical Association, Survey Research Methods Section, Anaheim, CA* (pp. 272-277). Alexandria, VA: American Statistical Association.

PURPOSE/OVERVIEW: Generalized variance functions often are employed when large numbers of estimates are to be published from a survey. Generalized variance functions lessen the volume of published reports where presentation of each standard error estimate would essentially double the size of the tabular presentations. In addition, generalized functions may facilitate secondary data analyses that were not conducted in the initial publications. Generalized variance functions also may provide more stable estimates of variance by diminishing the variability of the individual variance estimates themselves.

METHODS: In this paper, the authors studied some generalized models for the standard error of proportion estimates from the 1988 National Household Survey on Drug Abuse. A log-linear model based on the concept of a design effect was presented. The final model was evaluated against the simple average design effect (DEFF) model.

RESULTS/CONCLUSIONS: The authors concluded that, for these data, the domain-specific average DEFF model and the simple log-linear model both provided adequate generalized standard error models. They were surprised that the log-linear model including only effects for log(p), log(I-p), and log(n) performed so well. They expected that domain effects would be required in the model to account for the differential average cluster sizes by domain for this multistage sample design. It appeared that the slope for log(n) in the model being different from the simple random sampling (SRS) value of –0.5 accounted for most of the cluster size effect.

The feasibility of collecting drug abuse data by telephone

CITATION: Gfroerer, J. C., & Hughes, A. L. (1991). The feasibility of collecting drug abuse data by telephone. *Public Health Reports, 106*(4), 384-393.

PURPOSE/OVERVIEW: This article examines the feasibility of using telephone surveys to collect data on illicit drug use. It compares estimates on drug use from three data collection efforts, all carried out in 1988. The first, an in-person survey, was the 1988 National Household Survey on Drug Abuse (NHSDA). The other two surveys were telephone surveys, the Quick Response Survey (QRS) and the 1988 Texas Survey on Substance Use Among Adults.

METHODS: To assess the potential effects of noncoverage of nontelephone households on survey estimates, data from the 1988 NHSDA were used to compare estimates for those in telephone households with those in nontelephone households on lifetime and past year use of marijuana and cocaine. Estimates on illicit drug use and demographic items also were compared

between the 1988 NHSDA and the two telephone surveys. Comparable survey items were asked on all three surveys. For comparisons with the telephone surveys, NHSDA responses were edited to be consistent with questionnaire skip patterns in these surveys. For comparisons with estimates from the Texas Survey on Substance Use Among Adults, cases were restricted to the metropolitan areas of three certainty primary sampling units (PSUs) from the NHSDA. NHSDA data then were reweighted to match age group by metropolitan area totals from the Texas survey.

RESULTS/CONCLUSIONS: Within the NHSDA itself, drug use prevalence rates among those in telephone households were much lower than those in nontelephone households although most differences were not statistically significant. Estimates on drug use from QRS were consistently lower than those from NHSDA, even after controlling for differences in sample composition (QRS respondents reported higher levels of income and education than NHSDA respondents). For the combined metropolitan areas in Texas, one statistically significant difference was found between estimates from the NHSDA and the Texas survey on lifetime use of marijuana.

Improving the comprehension of reference periods

CITATION: Hubbard, M. L., Lessler, J. T., Graetz, K. A., & Forsyth, B. H. (1990). Improving the comprehension of reference periods. In *Proceedings of the 1990 American Statistical Association, Survey Research Methods Section, Anaheim, CA* (pp. 474-479). Alexandria, VA: American Statistical Association.

PURPOSE/OVERVIEW: This paper describes a pair of studies directed at examining alternative questioning strategies for the National Household Survey on Drug Abuse (NHSDA). The NHSDA provides comprehensive national data on drug use. As of 1990, its data were collected using a combination of interviewer- and self-administered questions. General questions on health and demographics were interviewer administered, while questions on more sensitive topics, such as the use of illegal drugs, employed self-administered answer sheets. Respondents frequently demonstrated a variety of errors when answering these types of questions. For example, in the 1988 survey, about 20 percent of the sample indicated they had used a drug at least once in the previous 12 months yet, in a subsequent question, indicated they had not used the drug in the previous 12 months.

METHODS: Under contract with the National Institute on Drug Abuse (NIDA), RTI designed a number of methodological studies exploring possible cognitive explanations and remedies for errors of this sort.

RESULTS/CONCLUSIONS: The results of these experiments were somewhat disappointing. Many surveys ask respondents to recall the number of events in specific reference periods. Because both bounded and intensive interviewing is expensive and time-consuming, it was hoped that these studies would demonstrate the effectiveness of less intensive techniques. This was not the case, although, in Study 2, the results were in the predicted direction. Perhaps respondents' estimates of the frequency of personal drug use were not generated by directly recalling individual incidences but by an estimation or averaging process that was not sensitive to the time periods used in these studies. If this was the case, it was possible that the ability to recall other types of events would be affected by using these types of less intensive procedures to

anchor the reference periods. However, it also may have been the case that either a more intensive interaction than was used in these studies was needed to anchor the reference periods or that respondents who were not in the anchoring condition were independently generating these kinds of anchors.

Analysis of survey data on drug experience by mode of data collection

CITATION: Hughes, A. L., & Gfroerer, J. C. (1990). Analysis of survey data on drug experience by mode of data collection. In *Proceedings of the 1990 American Statistical Association, Survey Research Methods Section, Anaheim, CA* (pp. 401-406). Alexandria, VA: American Statistical Association.

PURPOSE/OVERVIEW: In the late 1980s, the increasing concern over the drug abuse problem in the United States created a need for more data on the nature and extent of drug abuse. Policymakers demanded timely, accurate data at the national, State, and local levels to guide them in directing programs and funding toward the goal of reducing drug abuse and to measure progress in these programs. At the national level, the primary source of data on the prevalence of illicit drug use had been and remained the National Household Survey on Drug Abuse (NHSDA), sponsored by the National Institute on Drug Abuse (NIDA). The NHSDA was a probability sample survey of the U.S. household population that employed personal visit interviews with all selected respondents. The high cost of conducting household surveys raised interest in using telephone survey methodology to collect drug use prevalence data. In fact, several States had conducted drug use surveys by telephone and compared their State results with national NHSDA data (Frank, 1985; Spence, Fredlund, & Kavinsky, 1989). Some studies comparing data from telephone surveys to personal visit surveys showed that comparable health-related data (Massey, Barker, & Moss, 1979) and sociodemographic data (Monsees & Massey, 1979) could be obtained from the two methods. Other studies showed that sensitive data from telephone surveys were inferior to data collected from personal interviews (see Section 2 of this paper). However, very little research had been done investigating how the quality of data collected from drug use surveys differed by the two methods. The purpose of this paper was to present an analysis of the Quick Response Survey (QRS) and NHSDA in an attempt to evaluate the feasibility of using telephone survey methodology to collect data on illicit drug use.

METHODS: In 1988, NIDA funded a telephone survey on drug abuse through a cooperative agreement with the Food and Drug Administration (FDA) using a QRS contract. The QRS was conducted by Chilton Research Services at about the same time the 1988 NHSDA was in the field.

RESULTS/CONCLUSIONS: National estimates of illicit drug use from a telephone survey (QRS) and a personal visit survey (NHSDA) were compared. Overall, the results showed that estimates of illegal drug use from the QRS were significantly biased downward—a bias large enough to seriously distort confidence interval estimates. There were considerable differences in sample and drug use characteristics by phone status in the NHSDA data alone; however, because about 93 percent of the 1988 household population had phones, estimates from NHSDA telephone households were similar to estimates from the total sample. The effect of leading into sensitive questions on the use of illicit drugs by first asking about legal drugs cannot be ignored.

QRS respondents may have been more willing to admit their illegal drug use if they had been eased into these sensitive questions with less sensitive questions about such substances as cigarettes, alcohol, and prescription-type drugs, as is done on the NHSDA. This also could have had the negative effect of reduced response rates due to the length of the interview; however, although the authors concluded that this contextual issue needed further study, it was unlikely that this difference could have accounted for the substantial differences in reported drug use in the two surveys.

1992

Follow-up of nonrespondents in 1990

CITATION: Caspar, R. A. (1992). Follow-up of nonrespondents in 1990. In C. F. Turner, J. T. Lessler, & J. C. Gfroerer (Eds.), *Survey measurement of drug use: Methodological studies* (DHHS Publication No. ADM 92-1929, pp. 155-173). Rockville, MD: National Institute on Drug Abuse.

PURPOSE/OVERVIEW: This chapter presents the results of a special follow-up survey of persons who did not respond to the 1990 National Household Survey on Drug Abuse (NHSDA). The aim was to understand the reasons people chose not to participate, or were otherwise missed in the survey, and to use this information in assessing the extent of the bias, if any, that nonresponse introduced into the 1990 NHSDA estimates.

METHODS: To assess the impact of nonresponse, a follow-up study was undertaken on a subset of nonrespondents to the 1990 survey. The study was conducted in the Washington, DC, area, a region with a traditionally high nonresponse rate. The follow-up survey design included a $10 incentive and a shortened version of the instrument. The response rate for the follow-up survey was 38 percent.

RESULTS/CONCLUSIONS: The results of the follow-up study did not demonstrate definitively either the presence or absence of a serious nonresponse bias in the 1990 NHSDA. In terms of demographic characteristics, follow-up respondents appeared to be similar to the original NHSDA respondents. Estimates of drug use for follow-up respondents showed patterns that were similar to the regular NHSDA respondents. Only one statistically significant difference was found between the two groups for a composite measure of drug use (including cigarettes and alcohol) at any time during their lives, with follow-up respondents reporting higher rates than regular NHSDA respondents. Follow-up respondents also reported higher rates of lifetime use of cocaine, marijuana, alcohol, and cigarettes than regular NHSDA respondents, although the differences were not statistically significant. On the other hand, follow-up respondents had lower reported rates of past month use of cigarettes, alcohol, marijuana, and any drug use than regular NHSDA respondents, but these differences were not statistically significant. Another finding was that among those who participated in the follow-up survey, one third were judged by interviewers to have participated in the follow-up because they were unavailable for the main survey request. Finally, 27 percent were judged to have been swayed by the incentive and another 13 percent were judged to have participated in the follow-up due to the shorter instrument.

A follow-up study of nonrespondents to the 1990 National Household Survey on Drug Abuse

CITATION: Caspar, R. A. (1992). A follow-up study of nonrespondents to the 1990 National Household Survey on Drug Abuse. In *Proceedings of the 1992 American Statistical Association, Survey Research Section, Boston, MA* (pp. 476-481). Alexandria, VA: American Statistical Association.

PURPOSE/OVERVIEW: As of 1992, little was known about the drug use patterns of persons who were nonrespondents to the National Household Survey on Drug Abuse (NHSDA). Age, gender, race, and Hispanic origin were known from screening information for individual nonrespondents, but these data provided only minimal guidance in assessing the drug use patterns of persons who were not directly included in the regular survey estimates. To the extent that nonrespondents differed from respondents in their drug use and to the extent that NHSDA nonresponse adjustment procedures failed to take account of this difference, estimates from the NHSDA were subject to nonresponse bias. The issue of potential nonresponse bias was not a trivial one. The overall interview nonresponse rate in the 1990 NHSDA was 18 percent, with considerably higher rates in many locales. In the Washington, DC, metropolitan area, for example, the nonresponse rate was 27 percent.

METHODS: To assess the impact of such nonresponse, a follow-up study was undertaken of a subset of nonrespondents to the 1990 survey. For logistical reasons, the study was conducted in a single metropolitan area with a relatively high nonresponse rate. By offering nonrespondents a shortened questionnaire and monetary incentive, the authors hoped to persuade as many as possible to participate in the follow-up study. Their aim was to understand the reasons people chose not to participate—or were unavailable to participate in the survey—and to use this information in assessing the extent of the bias, if any, that nonresponse introduced into the 1990 NHSDA estimates.

RESULTS/CONCLUSIONS: The results of the follow-up study did not definitively demonstrate either the presence or the absence of a serious nonresponse bias in the 1990 NHSDA. For reasons of cost, the follow-up study was confined to the Washington, DC, metropolitan area, and the results may not have been generalizable to other areas of the country. Similarly, with a response rate of 38 percent, there remained a sizable majority of sample nonrespondents for whom no information was obtained. Anecdotal information from follow-up interviewers suggested that these hard-core nonrespondents may have differed significantly in their drug use behaviors from persons interviewed as part of either the regular NHSDA or the follow-up study. In terms of demographic characteristics, follow-up respondents appeared to be similar to the original NHSDA respondents. Estimates of drug use for follow-up respondents showed patterns similar to the regular NHSDA respondents. Only one statistically significant difference was found between the two groups—for the composite measure of drug use at anytime during their lives.

From both the qualitative and the quantitative data presented here, it would appear that the NHSDA nonrespondents who were interviewed in the follow-up study were quite similar to the respondents interviewed as part of the regular NHSDA data collection. Interviewers working on the follow-up noted that persons who continued to refuse to participate appeared to have

something to hide and to be afraid of answering questions about drugs. Whether this was indicative of higher rates of drug use among hard-core nonrespondents was unknown. Adding the follow-up cases to the regular NHSDA sample made little difference to the NHSDA estimates of drug use prevalence. The authors did not know, however, how convincing these remaining hard-core nonrespondents would affect NHSDA's estimates. Because the authors could not follow up all nonrespondents, the assessment of nonresponse bias provided by this method was, of necessity, incomplete. Nevertheless, it could indicate the potential impact on NHSDA prevalence estimates of alternative survey procedures (e.g., selective use of monetary incentives) to increase response rates.

Inconsistent reporting of drug use in 1988

CITATION: Cox, B. G., Witt, M. B., Traccarella, M. A., & Perez-Michael, A. M. (1992). Inconsistent reporting of drug use in 1988. In C. F. Turner, J. T. Lessler, & J. C. Gfroerer (Eds.), *Survey measurement of drug use: Methodological studies* (DHHS Publication No. ADM 92-1929, pp. 109-153). Rockville, MD: National Institute on Drug Abuse.

PURPOSE/OVERVIEW: This chapter describes the inconsistencies that occurred in the reporting of drug use in the 1988 National Household Survey on Drug Abuse (NHSDA). A methodological study was conducted to locate and measure the faulty data in the 1988 NHSDA; results are summarized in this chapter. The study examined inconsistent responses within and among various sections of the questionnaire used for the 1988 NHSDA.

METHODS: For this analysis, the authors used data taken from a partially edited data file. Many of the inconsistent responses included in the analyses reported in this chapter could be eliminated by logical editing of the data. This methodological analysis was designed to identify those aspects of the survey questionnaire and survey process that caused difficulty for participants. It was determined that these aspects of the survey, in turn, should become primary targets for revision in future administrations of the NHSDA.

RESULTS/CONCLUSIONS: The first section of this chapter provides an overall indication of approximately how many respondents made errors, where the errors were made, and the demographic distribution of the respondents who made the errors. Subsequent sections present further details in the inconsistencies in reporting found on the alcohol, marijuana, and cocaine forms, and in reporting concomitant drug use. Findings include the following: (1) A total of 24 percent of those who reported using at least one drug in their lifetime answered at least one question on drug use inconsistently. Among respondents who reported using at least one drug in the past 12 months, 31 percent answered at least one question on drug use inconsistently. (2) Inconsistencies occurred most frequently on the alcohol form, although on a percentage basis, inconsistencies in alcohol reporting occurred at lower rates than for many less frequently used substances. (3) No inconsistencies occurred within the cigarette form, which, unlike other forms, was administered by the interviewer. The same was true for the sedative, tranquilizer, stimulant, and analgesic forms, which provided an option to skip the form if they had never used these drugs. (4) Young respondents, those between the ages of 12 to 17, were less likely to provide inconsistent responses than older respondents.

A method for identifying cognitive properties of survey items

CITATION: Forsyth, B. H., & Hubbard, M. L. (1992). A method for identifying cognitive properties of survey items. In *Proceedings of the 1992 American Statistical Association, Survey Research Methods Section, Boston, MA* (pp. 470-475). Alexandria, VA: American Statistical Association.

PURPOSE/OVERVIEW: Question wording can lead to error and bias survey measurement (e.g., Forsyth & Lessler, 1991a; Groves, 1989; Sudman & Bradburn, 1974; Turner & Marlin, 1984). This paper reports on research to develop and test a method for identifying survey items that are difficult for respondents to answer due to their cognitive demands. For example, items may be difficult to answer if the wording is difficult to understand, if response requires detailed memory recall, or if response categories fail to cover the range of respondent experience. The authors' aim was to design a taxonomy of item characteristics that could be used to identify potentially problematic survey items. The research reported was part of the methodological study of the National Household Survey on Drug Abuse (NHSDA) sponsored by the National Institute on Drug Abuse (NIDA). Major goals of the larger research project were to (1) identify potential sources of measurement error in NHSDA, (2) revise survey materials and survey procedures that seem to contribute to avoidable measurement error, and (3) test revisions and identify improved strategies for measuring patterns drug use. This paper focuses on reducing measurement errors that arise from the cognitive processes that respondents use when answering survey items.

METHODS: The appraisal results summarized here were used as one basis for identifying method improvements to test under more formal laboratory and field test conditions. Based in part on these appraisal results, the authors developed three sets of improvements. First, they used laboratory and field test procedures to investigate test decomposition approaches for defining technical terminology and complex reference sets. Second, they used laboratory and field test methods to test procedures for anchoring reference periods. Third, they used field test methods to test experimental questionnaire materials that eliminated hidden questions by using branching instructions and skip patterns.

RESULTS/CONCLUSIONS: As reported in other papers in the 1992 volume, the experimental and field test results suggested that the authors' appraisal methodology made an important contribution to identifying sources of response inconsistencies, response biases, and response variability. Additional research is necessary before this coding scheme can be used as a general purpose tool for analyzing survey items. The authors have worked to clarify, refine, and trim their coding categories, collapsing some while expanding upon others. In addition, they indicate a need for further research to provide valid tests of the coding scheme once it has been refined. Although further development and testing is necessary, the authors believe that they have begun to develop a cost-effective method for systematizing expert evaluations and for identifying and cataloging critical aspects of questionnaire wording and format.

An overview of the National Household Survey on Drug Abuse and related methodological research

CITATION: Gfroerer, J. (1992). An overview of the National Household Survey on Drug Abuse and related methodological research. In *Proceedings of the 1992 American Statistical Association, Survey Research Methods Section, Boston, MA* (pp. 464-469). Alexandria, VA: American Statistical Association.

PURPOSE/OVERVIEW: This paper provides a brief description of the history of the National Household Survey on Drug Abuse (NHSDA) and a summary of methodological research carried out prior to 1992.

METHODS: A summary of the methodological research that had been done in conjunction with NHSDA is given.

RESULTS/CONCLUSIONS: The results summarized in this paper include the following: (1) A field test conducted in 1990 found lower rates of drug use reporting for an interviewer-administered survey than self-administered surveys, particularly for more recent use. The magnitudes of the differences varied by substance reported (smallest for alcohol, larger for marijuana and even larger for cocaine). Differences in estimates between the interviewer- and self-administered forms were larger for youths than adults. (2) A skip pattern experiment conducted in 1992 found that prevalence rates for marijuana and cocaine were lower in a version that employed skip patterns than the main NHSDA sample (which did not employ skip patterns). (3) A follow-up survey of nonrespondents in the Washington, DC, area for the 1990 NHSDA had a 38 percent response rate. The most frequent reason offered for participating in the follow-up survey was that the respondent was unavailable for the initial survey request. (4) A validation study in which a sample of clients recently discharged from drug treatment facilities were administered the NHSDA interview found higher levels of underreporting on the survey (relative to information from treatment records) for more socially deviant behaviors (e.g., higher levels of underreporting for heroin and cocaine, lower levels of underreporting for marijuana). The study also found very low levels of reporting drug treatment on the interview (only 38 percent). (5) Finally, additional research summarized here found that reporting of drug use by youths is more likely under conditions of complete privacy.

The incidence of illicit drug use in the United States, 1962–1989

CITATION: Gfroerer, J., & Brodsky, M. (1992). The incidence of illicit drug use in the United States, 1962-1989. *British Journal of Addiction, 87*(9), 1345-1351.

PURPOSE/OVERVIEW: As of the early 1990s, epidemiological drug studies over the prior four decades had primarily focused on the prevalence of illicit drug use and not on the incidence of use as well as trends. In this paper, retrospective drug use data from the National Household Survey on Drug Abuse (NHSDA) were used to provide estimates for the number of new drug users each year from 1962 to 1989.

METHODS: Data from 5 years of the NHSDA were combined to yield a total sample size of 58,705 respondents. The variables of interest in the surveys were date of birth, date of interview, and reported age at first use for each drug (heroin, cocaine, marijuana, and hallucinogens), which allowed the researchers to pinpoint an exact date of first drug use for each respondent. The methods and validity of this approach also were evaluated.

RESULTS/CONCLUSIONS: The findings show that the peak year for new marijuana users was in 1973 and remained roughly stable until about 1980, at which point it started to drop. The average age of new users remained around the late teens since 1962. Patterns for hallucinogen use were similar to marijuana use. Peak estimates for incidence of cocaine use occurred in 1982 and subsequently had seen a dropoff. Retrospective estimates for drug use incidence have some limitations in that they exclude drug users who died before 1985, which could be significant for the estimates prior to 1962. Retrospective reports also are affected by recall bias in that estimates for the early years relied on a recall of over 20 years. However, advantages of the retrospective estimates were that a large sample size was able to provide stability in the estimates.

Introduction. In C. F. Turner, J. T. Lessler, & J. C. Gfroerer (Eds.), Survey measurement of drug use: Methodological studies

CITATION: Gfroerer, J. C., Gustin, J., & Turner, C. F. (1992). Introduction. In C. F. Turner, J. T. Lessler, & J. C. Gfroerer (Eds.), *Survey measurement of drug use: Methodological studies* (DHHS Publication No. ADM 92-1929, pp. 3-10). Rockville, MD: National Institute on Drug Abuse.

PURPOSE/OVERVIEW: This is the introduction to a monograph that presents a range of studies assessing the accuracy of alternative methods for survey measurement of drug use. This volume reports the results of a program of methodological research designed to evaluate and improve the accuracy of measurements made in the National Household Survey on Drug Abuse (NHSDA), the Nation's major survey for monitoring drug use. Most of the research reported in this volume began in 1989 and was conducted by a team of scientists from the Research Triangle Institute and the National Institute on Drug Abuse.

METHODS: This introductory chapter describes the research components and briefly outlines the origins and purposes of the NHSDA and its methodological research program. The final section of this chapter describes the organization of the volume and the contents of the chapters.

RESULTS/CONCLUSIONS: N/A.

Collecting data on illicit drug use by phone

CITATION: Gfroerer, J. C., & Hughes, A. L. (1992). Collecting data on illicit drug use by phone. In C. F. Turner, J. T. Lessler, & J. C. Gfroerer (Eds.), *Survey measurement of drug use: Methodological studies* (DHHS Publication No. ADM 92-1929, pp. 277-295). Rockville, MD: National Institute on Drug Abuse.

PURPOSE/OVERVIEW: This chapter evaluates the feasibility of using telephone surveys to collect data on illicit drug use. It compares estimates on drug use from three data collection efforts, all carried out in 1988: (1) the 1988 National Household Survey on Drug Abuse (NHSDA), (2) the Quick Response Survey (QRS), and (3) the 1988 Texas Survey on Substance Use Among Adults.

METHODS: Using data from the 1988 NHSDA, estimates for those in telephone households were compared with those in nontelephone households on lifetime and past year use of marijuana and cocaine in order to examine the effects of noncoverage of nontelephone households in a telephone survey. Estimates on illicit drug use and demographic items from the 1988 NHSDA also were compared with those from QRS, after NHSDA data were reedited for consistency with QRS skip patterns. Finally, estimates from the 1988 NHSDA from Texas were compared with those from the Texas Survey on Substance Use Among Adults. For this analysis, data were restricted to the metropolitan areas of three certainty primary sampling units (PSUs) from NHSDA. NHSDA estimates were combined with those from the QRS and then these in turn were compared with estimates from the Texas survey.

RESULTS/CONCLUSIONS: Although some differences were not statistically significant, drug use prevalence among telephone households was generally much lower than that among nontelephone households in NHSDA. Estimates on drug use from QRS were consistently lower than those from NHSDA, even after accounting for differences in sample composition (QRS respondents reported higher levels of income and education than NHSDA respondents). One statistically significant difference was found (among four items examined) between estimates from NHSDA for the combined metropolitan areas in Texas and the Texas survey (for lifetime use of marijuana).

The intersection of drug use and criminal behavior: Results from the National Household Survey on Drug Abuse

CITATION: Harrison, L., & Gfroerer, J. (1992). The intersection of drug use and criminal behavior: Results from the National Household Survey on Drug Abuse. *Crime and Delinquency, 38*(4), 422-443.

PURPOSE/OVERVIEW: The purpose of this paper is to explore the relationship between drug use and criminal behavior focusing on the difference between licit and illicit drug use.

METHODS: In 1991, questions about criminal involvement and criminal behavior were added to the National Household Survey on Drug Abuse (NHSDA). A total of 32,594 respondents completed the NHSDA interview, which contained self-report questions on drug use and criminality. The relationship between drug use and crime was analyzed using cross-tabulations and logistic regression.

RESULTS/CONCLUSIONS: The results showed a positive correlation between criminal behavior and higher levels of drug use. Respondents who did not report using drugs or alcohol during the past year had the lowest levels of criminal behavior. After controlling for sociodemographic characteristics, such as age, gender, and race, higher drug use still was

associated significantly with criminal behavior. The results showed further findings by the type of criminal behavior and other individual-level covariates.

Effects of decomposition of complex concepts

CITATION: Hubbard, M. L., Pantula, J., & Lessler, J. T. (1992). Effects of decomposition of complex concepts. In C. F. Turner, J. T. Lessler, & J. C. Gfroerer (Eds.), *Survey measurement of drug use: Methodological studies* (DHHS Publication No. ADM 92-1929, pp. 245-264). Rockville, MD: National Institute on Drug Abuse.

PURPOSE/OVERVIEW: This chapter reports the results of the authors' experiment and compares the responses obtained using an alternative measurement strategy with those obtained with the National Household Survey on Drug Abuse (NHSDA) in the early 1990s.

METHODS: In the experiment, the authors tested a questionnaire in which complex concepts were decomposed into a number of simpler elements for which more straightforward questions could be formulated.

RESULTS/CONCLUSIONS: The evidence with regard to measurements of the nonmedical use of psychotherapeutics indicated that decomposing the NHSDA question on this issue substantially increased the reporting of such use. The results were striking, particularly for painkillers. The estimated prevalence of nonmedical use of painkillers more than doubled when the authors decomposed the concept into constituent parts and asked respondents separate questions about each part. For stimulants and tranquilizers, the differences were not as striking as for painkillers, but they were substantial. The analysis of responses to the identical question using this new questioning strategy suggested that some respondents may have used a personal definition of nonmedical use regardless of the instructions provided on the questionnaire. It became clear to the authors that survey designers cannot rely on respondents to follow instructions to classify usage as medical or nonmedical.

1992 National Household Survey on Drug Abuse: Findings of first quarter skip test: Final report

CITATION: Lessler, J. T., & Durante, R. C. (1992, October 7). *1992 National Household Survey on Drug Abuse: Findings of first quarter skip test: Final report* (prepared for the Substance Abuse and Mental Health Services Administration, Contract No. 271-91-5402, RTI/5071). Research Triangle Park, NC: Research Triangle Institute.

PURPOSE/OVERVIEW: In 1990, the National Institute on Drug Abuse (NIDA) conducted a large methodological field test of the National Household Survey on Drug Abuse (NHSDA), primarily to evaluate the effect of using interviewer-administered versus self-administered questionnaires. In this study, some questionnaire answer sheets incorporated skip patterns. It was found that, in general, respondents were able to follow skip patterns that were not too complex, particularly if skips were always to the top of a page. Generally, when errors occurred, they resulted in respondents answering additional questions unnecessarily, so there was no loss of

data. However, the design of the field test did not allow a determination of the effect of skip patterns on reporting of drug use.

In an attempt to address this, NIDA conducted an experiment during the first 3 months of 1992 to test a new questionnaire that incorporated skip patterns into the drug use answer sheets. An experimental questionnaire was developed that included a number of skip patterns that allowed respondents to skip out of entire sections of questions if they responded "no" to an initial question on whether they had used a drug. This questionnaire was called the skip version. The regular NHSDA questionnaire was called the nonskip version. Differences between the two questionnaires varied by section in the questionnaire.

METHODS: The Skip Pattern Experiment was embedded in the first quarter 1992 NHSDA sample. One eighth of the first quarter sample was randomly assigned to receive the skip questionnaire, while the other seven eighths received the nonskip version. Assignment of questionnaire versions to sampled dwelling units was done in advance of any contact by field staff. Allocation of the skip version was done within sample segments to maximize the power of statistical comparisons between the two groups. Interviewers were trained to administer both versions of the questionnaire. Overall, the nonskip version was administered to 7,149 respondents, and the skip version was administered to 974 respondents.

RESULTS/CONCLUSIONS: Overall, this methodological study indicated that using skip patterns tended to reduce the prevalence of illicit drug use. There also was an indication that the bias due to using skips would not be uniform across different populations, as it seemed to be more pronounced among respondents with more education. It was not possible to conclude whether the lack of privacy or the desire to avoid additional questions was operating. However, it is interesting to note that the skip in the alcohol questions had no apparent impact. This would suggest that the sensitivity of the illicit drug use questions and the loss of privacy in the skip version are most important.

Effects of mode of administration and wording on reporting of drug use

CITATION: Turner, C. F., Lessler, J. T., & Devore, J. W. (1992). Effects of mode of administration and wording on reporting of drug use. In C. F. Turner, J. T. Lessler, & J. C. Gfroerer (Eds.), *Survey measurement of drug use: Methodological studies* (DHHS Publication No. ADM 92-1929, pp. 177-219). Rockville, MD: National Institute on Drug Abuse.

PURPOSE/OVERVIEW: This chapter assesses the impact of question wording and mode of administration on estimates of prevalence of drug use from the National Household Survey on Drug Abuse (NHSDA).

METHODS: The authors describe the design of the methodological field test conducted in 1990 and review estimates of the prevalence of self-reported drug use obtained by the different questionnaire versions used in this field test.

RESULTS/CONCLUSIONS: Although the results were not always consistent for every substance examined, on balance, the results indicate that having interviewers administer the

questionnaire reduces the reporting of drug use. This conclusion is supported by the finding that lack of privacy during an interview had a negative effect on the reporting of drug use, particularly for respondents 12 to 17 years of age for whom the person present is likely to be a parent. The authors' finding that self-administered forms yield more reports of drug use does not appear to be due to a greater number of marking errors.

Effects of mode of administration and wording on data quality

CITATION: Turner, C. F., Lessler, J. T., George, B. J., Hubbard, M. L., & Witt, M. B. (1992). Effects of mode of administration and wording on data quality. In C. F. Turner, J. T. Lessler, & J. C. Gfroerer (Eds.), *Survey measurement of drug use: Methodological studies* (DHHS Publication No. ADM 92-1929, pp. 221-243). Rockville, MD: National Institute on Drug Abuse.

PURPOSE/OVERVIEW: In this 1992 chapter, the authors review the impact of various factors on the completeness and consistency of the reporting of drug use on the National Household Survey on Drug Abuse (NHSDA).

METHODS: The authors summarize findings of their investigations of selected aspects of the quality of the data produced by the different versions of the survey questionnaire and the different modes of administration used in a 1990 NHSDA field test. The authors consider the rates of nonresponse, the extent to which respondents and interviewers correctly executed the branching instructions embedded in the questionnaires, and the internal consistency of the reports of drug use given by respondents.

RESULTS/CONCLUSIONS: The results of the authors' analyses provide the key ingredients for conclusions concerning the relative quality of the data provided by the different versions of the survey evaluated in the 1990 NHSDA field test. Respondents were, in general, capable of responding to a self-administered form, even when that form included many branching or skip instructions. The changes made to the NHSDA questionnaire generally improved respondent understanding of the questions and thereby improved the quality of the data collected.

Survey measurement of drug use: Methodological studies

CITATION: Turner, C. F., Lessler, J. T., & Gfroerer, J. C. (Eds.). (1992). *Survey measurement of drug use: Methodological studies* (DHHS Publication No. ADM 92-1929). Rockville, MD: National Institute on Drug Abuse.

PURPOSE/OVERVIEW: This 1992 volume presents a range of studies assessing the accuracy of alternative methods for the survey measurement of drug use. It contains demonstrations of the liability of such measurements in response to variation in the measurement procedures that are employed. The editors urge that these demonstrations should not be taken as an indictment of the use of surveys to measure drug use. The data obtained from a measurement process are properly seen to be a joint function of the phenomenon under study and the protocol used to make the measurement. This joint parentage is reflected in the frequent practice of reporting such

measurements by reference to both the measurement outcome and the protocol used in the measurement.

METHODS: The volume is divided into six parts, the first of which is the introduction. Part II contains two chapters that present the results of cognitive research on the National Household Survey on Drug Abuse (NHSDA) questionnaire. Part III comprises three chapters that analyze different aspects of past NHSDAs to identify problems. Part IV reports the results of a large-scale field experiment that tested the new version of the NHSDA questionnaire. Part V presents the results of two studies that complement the field experiment. Part VI is the single concluding chapter that summarizes major findings.

RESULTS/CONCLUSIONS: The various chapters in this volume demonstrate the appropriateness of such practices in the reporting and use of survey measurements of drug use. The authors offer a wide-ranging view of the impact of measurement procedures on survey measurements of drug use, and they introduce new techniques for diagnosing problems with survey questionnaires and for designing improved ones.

Item nonresponse in 1988

CITATION: Witt, M. B., Pantula, J., Folsom, R. E., & Cox, B. G. (1992). Item nonresponse in 1988. In C. F. Turner, J. T. Lessler, & J. C. Gfroerer (Eds.), *Survey measurement of drug use: Methodological studies* (DHHS Publication No. ADM 92-1929, pp. 85-108). Rockville, MD: National Institute on Drug Abuse.

PURPOSE/OVERVIEW: This chapter examines the patterns of missing data in the 1988 National Household Survey on Drug Abuse (NHSDA). The quality of a survey instrument may be gauged in part by the completeness and internal consistency of the data it produces. Survey instruments that produce more and more consistent measurements are preferable. Evidence of incomplete or inconsistent responses to a survey can be used to identify the aspects of a survey that are in need of remediation. Respondents' failure to answer questions may, for example, reflect their misunderstanding of the survey instructions or bafflement with a particular question.

METHODS: The authors assessed the levels and patterns of nonresponse to questions in the 1988 NHSDA. The nonresponse was not strongly concentrated on any one form, nor focused on any one type of question. Nevertheless, a few general statements were made about the nonresponse. A disproportionately large number of respondents were people who never used the drug that was discussed on the form. A great deal of nonresponse resulted from those questions that asked about an individual's concomitant drug use and those questions that asked the age at which a respondent first took a particular drug.

RESULTS/CONCLUSIONS: The authors concluded that the high rate of nonresponse to questions regarding age occurred because people had a hard time recalling what had happened in the past. The substantial nonresponse on concomitant drug use may be due to the question's sensitivity.

1993

Self-reported drug use data: What do they reveal?

CITATION: Harrison, E. R., Haaga, J., & Richards, T. (1993). Self-reported drug use data: What do they reveal? *American Journal of Drug and Alcohol Abuse, 19*(4), 423-441.

PURPOSE/OVERVIEW: The purpose of this article was to measure the accuracy of self-reported drug use by examining the consistency of the self-reports over time.

METHODS: Using data from the 1985 and 1990 National Household Surveys on Drug Abuse (NHSDAs), trends in self-reported drug use for marijuana and cocaine were compared across time. Two years of data from the National Longitudinal Survey Youth Cohort (NLS-Y) also were analyzed to observe any discrepancies in reported drug use over time.

RESULTS/CONCLUSIONS: The results from the NHSDA analysis revealed a declining trend in drug use starting in the late 1980s, which was consistent with data collected from other self-report studies. A cohort analysis from the NHSDA data revealed that fewer people in each birth cohort reported ever using marijuana or cocaine in 1990 than in 1985. This phenomenon occurred particularly in the birth cohorts born in the mid-1950s to the mid-1960s. The decrease in reported lifetime drug use over time led researchers to conclude that either people were not as willing to report drug use as they used to be or else people's perceptions of what constitutes drug use had changed. Results from the NLS-Y data were consistent with the NSDUH results in that 20 percent of the respondents in the 1984 survey who admitted to ever using drugs denied using them in the 1988 survey. The most significant indicators of discrepancy in reported drug use were being African American and the frequency of drug use. These results raised the question of whether outside factors, such as public disapproval or drug use, affect respondents' willingness to reveal their prior drug use.

Report on 1990 NHSDA-Census Match

CITATION: Parsley, T. (1993). *Report on 1990 NHSDA-Census Match* (RTI/4469-15, prepared for the Substance Abuse and Mental Health Services Administration, Contract No. 271-89-8333). Research Triangle Park, NC: Research Triangle Institute.

PURPOSE/OVERVIEW: In 1991, the National Institute on Drug Abuse (NIDA) and the Research Triangle Institute (RTI) cooperated with the U.S. Bureau of the Census on a study of nonresponse in surveys. This effort was part of the multiagency work conducted by Robert M. Groves and Mick P. Couper at the Census Bureau to study nonresponse in seven major Federal Government surveys, including the National Household Survey on Drug Abuse (NHSDA). The study involved linking data from a sample of respondents and nonrespondents to the 1990 NHSDA with their 1990 decennial census data to provide descriptive information about NHSDA nonrespondents. NHSDA and census records were linked by Census Bureau clerks using primarily address and other location information.

METHODS: A total of 5,030 NHSDA households were selected to be matched to 1990 census records. All 860 screener nonresponse households were selected. In addition, all 1,821 households with at least one interview nonrespondent were selected for matching. To allow comparisons between respondents and nonrespondents on the census items, a random systematic sample of 1,938 households was selected in which all sample persons completed the interview. Finally, to assess the accuracy of interviewers' classifications, a sample of 411 cases classified as vacant, temporary dwellings, or nonhousing units by the NHSDA interviewer was selected.

RESULTS/CONCLUSIONS: The Census Match Study demonstrates that response rates are not constant across various interviewer, respondent, household, and neighborhood characteristics. To the extent that rates of drug use vary by these same characteristics, bias due to nonresponse may be a problem. However, it is not always the case that low response rates occur in conjunction with high drug use prevalence. Some populations with low response rates (e.g., older adults and high-income populations) tend to have low rates of drug use. On the other hand, some populations (e.g., residents of large metropolitan areas and men) have low response rates and high drug use rates. In estimating overall prevalence, many of these potential sources of bias would be in opposite directions and would therefore tend to cancel each other.

Evaluation of results from the 1992 National Household Survey on Drug Abuse (NHSDA)

CITATION: Peer Review Committee on National Household Survey on Drug Abuse (NHSDA) [James T. Massey, Chair; Marc Brodsky; Tom Harford; Lana Harrison; Dale Hitchcock; Ron Manderscheid; Nancy Pearce; Marilyn Henderson; Beatrice Rouse; & Ron Wilson]. (1993, June 3). *Evaluation of results from the 1992 National Household Survey on Drug Abuse (NHSDA)* [memo analyzing the decrease in drug prevalence among the black population between the 1992 NHSDA and previous NHSDAs]. Research Triangle Park, NC: Research Triangle Institute.

PURPOSE/OVERVIEW: The preliminary results from the first two quarters of 1992 raised some concerns at the Research Triangle Institute (RTI) and the Substance Abuse and Mental Health Services Administration (SAMHSA) because of some differences between the 1992 results and the survey results from 1988, 1990, and 1991. The main concern was a greater than expected decline in reported drug use in 1992, primarily among blacks. Particularly puzzling was a drop in the reported lifetime use of illicit drugs, which would not be expected to change from one year to the next. Because of the survey differences between years, the Office of Applied Studies (OAS) formed a Peer Review Committee (PRC) to evaluate the results from the 1992 National Household Survey on Drug Abuse (NHSDA) and to make recommendations about the release and publication of the results.

METHODS: The first step in the evaluation process was to review the key results from the 1988 to 1992 NHSDAs and to identify the key differences and any inconsistencies in trends from 1988 through 1992. The committee reached the same conclusions that had been reached by the OAS and RTI staffs. Namely, the 1992-reported drug use for blacks was consistently lower than the 1991 results. A number of possible causes for the 1992 decline in drug use were explored by

analyzing the results from the 1988 to 1992 surveys, although some of the possible explanations could be evaluated only indirectly or subjectively.

RESULTS/CONCLUSIONS: The consensus of the committee was that the observed differences between 1991 and 1992 cannot be explained by a single factor, although several small differences were found among the factors examined. In conducting its analysis of the NHSDA, the committee concluded that the design and procedures for sampling, weighting, editing, and imputing the survey results were statistically sound. Great care had been taken by the National Institute on Drug Abuse (NIDA), OAS, and RTI to implement the survey procedures and to evaluate the quality of the results. The committee concluded that the unexpected decrease in lifetime drug use among blacks was an example of what can occasionally occur in survey estimates, particularly when a large number of different estimates are generated and comparisons are made. Often, a review of the procedures will uncover an error in the process. In other cases, such as the NHSDA, an explanation for the unexpected results may never be found.

Drug use measurement: Strengths, limitations, and recommendations for improvement

CITATION: U.S. General Accounting Office. (1993). *Drug use measurement: Strengths, limitations, and recommendations for improvement* (GAO/PEMD-93-18, Report to the Chairman, Committee on Government Operations, House of Representatives, Program Evaluation and Methodology Division). Washington, DC: U.S. Government Printing Office.

PURPOSE/OVERVIEW: Policymakers, researchers, and planners must have accurate drug use information if they are to properly assess the Nation's current drug prevalence pattern and trends. However, the quality of data can be constrained by methodological problems, available research technology, and environmental and budgetary limitations. Because millions of dollars are spent on drug prevalence studies, it is important to evaluate the current state of drug use measurement practices within a cost feasibility context. In response to a request by the Chairman of the House Committee on Government Operations, the GAO investigated the issue of drug use measurement by (1) reporting on three nationally prominent drug studies, (2) assessing the methodological strengths and limitations of each of these studies, and (3) developing recommendations for the improvement of drug prevalence estimates.

METHODS: The GAO examined the National Household Survey on Drug Abuse (NHSDA), the High School Senior Survey (HSSS), and the Drug Use Forecasting (DUF) study of booked arrestees. The GAO evaluated the methodological strengths and limitations of these three studies in terms of the degree to which their research operations satisfied generally accepted criteria. The GAO also developed guidelines for improving drug prevalence estimates, focusing particularly on high-risk groups.

RESULTS/CONCLUSIONS: NHSDA is a sophisticated study of drug use patterns and trends, but is limited by the exclusion of groups at high risk for drug use, problematic measurement of heroin and cocaine use, and reliance on subject self-reports. HSSS has similar limitations. Therefore, both surveys provide conservative estimates of drug use. DUF cannot be generalized to booked arrestees in the geographic areas sampled. The GAO found that drug prevalence

estimates could be improved and money could be saved if NHSDA and HSSS were administered in alternate years. The GAO cited several ways of validating these two surveys and estimating the extent of underreporting. Expanding the subsamples of current surveys and conducting new studies aimed at hard-to-reach, high-risk groups should improve the coverage of underrepresented target populations.

1994

Repeated measures of estimation of measurement bias for self-reported drug use with applications to the National Household Survey on Drug Abuse

CITATION: Biemer, P., & Witt, M. (1994). Repeated measures of estimation of measurement bias for self-reported drug use with applications to the National Household Survey on Drug Abuse. In *Proceedings of the National Institute on Drug Abuse technical review: The validity of self-reported drug use*. Rockville, MD: National Institute on Drug Abuse.

PURPOSE/OVERVIEW: Direct estimates of response bias in self-reports of drug use in surveys require that essentially error-free determinations of drug use be obtained for a subsample of survey respondents. The difficulty of obtaining determinations that are accurate enough for estimating validity is well documented in the literature. Methods such as specimen (hair, urine, etc.) analysis, proxy reports, and the use of highly private and anonymous modes of interview all have to contend with error rates that may be only marginally lower than those of the parent survey. Thus, any methodology for direct validity estimation must rely to some extent on approximations and questionable assumptions.

METHODS: In this article, the authors considered a number of methods that rely solely on repeated measures data to assess response bias. Because the assumptions associated with these approaches do not require highly accurate second determinations, they may be more easily satisfied in practice. One such method for bias estimation for dichotomous variables that is considered in some detail provides estimates of misclassification probabilities in the initial measurement without requiring that the second measure be accurate or even better than the first. This methodology does require, however, that two subpopulations exist that have different rates of prevalence but whose probabilities of false-positive and false-negative error are the same. The applicability of these methods for self-reported drug use is described and illustrated using data from the National Household Survey on Drug Abuse (NHSDA). In the discussion of the results, the importance of these methods for assessing the validity of self-reported drug use is examined.

RESULTS/CONCLUSIONS: In this paper, a general model for studying misclassification in self-reported drug use was presented, and the model then was extended to the case where two measurements of the same characteristic are available for a sample of respondents. For the two measurements case, the general model required seven parameters while only three degrees of freedom were available for estimation. It is shown how the assumptions typically made for test-retest, true value, improved value, and Hui and Walter methods relate to the general model.

Estimating substance abuse treatment need from the NHSDA

CITATION: Epstein, J. F., & Gfroerer, J. C. (1994). Estimating substance abuse treatment need from the NHSDA. In *Proceedings of the 1994 Joint Statistical Meetings, American Statistical Association, Survey Research Methods Section, Toronto, Ontario, Canada* (pp. 559-564). Alexandria, VA: American Statistical Association.

PURPOSE/OVERVIEW: By the early 1990s, several methods had been developed to estimate treatment need using the National Household Survey on Drug Abuse (NHSDA). These methods were generally developed to estimate drug abuse treatment need, not alcohol treatment need. The National Institute on Drug Abuse (NIDA) developed an illicit drug index in 1989 that defined heavy drug users who need treatment as those who had used illicit drugs at least 200 times in the past year. An estimated 4 million Americans met this criteria in 1988. Another model developed by NIDA was based on reported problems and symptoms of abuse or dependence on illicit drugs. This method was an approximation of criteria in the 3rd edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-III) published by the American Psychiatric Association (APA, 1980). This method produced an estimate of 3.5 million people with drug abuse or dependence in 1991. A study sponsored by the Institute of Medicine (IOM) in 1990 developed a methodology for estimating drug abuse treatment need using a combination of frequency of use and problems/symptoms indicators from the NHSDA. The IOM also supplemented these NHSDA estimates with data from prison and homeless populations to produce a more comprehensive estimate of need. This method resulted in an estimated 4.6 million people needing treatment in 1988. Throughout this paper, this method is referred to as the FREQ method. As of 1994, a method for estimating treatment need from the NHSDA had been considered by the Office of Applied Studies (OAS) of the Substance Abuse and Mental Health Services Administration (SAMHSA). This method used an algorithm that approximated the criteria from the 3rd revised edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-III-R) for illicit drug dependence (APA, 1987). This method is referred to as the DEP method.

METHODS: This paper compared the FREQ method and the DEP method. Because the FREQ method did not apply to alcohol, alcohol was not included in the comparison. One objective of this comparison was to better understand the strengths and weaknesses of these two methods, providing a basis for further improvements in treatment need estimation by SAMHSA. This comparison is based on an analysis of the 1991 NHSDA.

RESULTS/CONCLUSIONS: From the investigation of the DEP and FREQ estimates of treatment need on different populations, the following conclusions were reached: (1) Many populations of serious drug users were only partially assigned as needing treatment by both the FREQ and the DEP methods. For example, less than half of the persons who used needles in the past year, less than half of the persons who used heroin in the past year, and less than half of the persons who used cocaine weekly or more often were estimated to need treatment by each method. (2) Both methods estimated a very small proportion of persons using prescription drugs as needing treatment. (3) The two methods did not estimate the same people as needing treatment (kappa statistics were less than .5 for all populations and less than .2 for more than half the populations). (4) The FREQ method generally estimated a larger proportion of persons as needing treatment. In summary, the FREQ method defined a broader group of drug abusers as needing treatment, while the DEP method defined a smaller more severely impaired group of drug abusers as needing treatment. For some populations clearly in need of treatment, the FREQ estimate was better than the DEP estimate because it estimated a larger proportion of the persons in these populations as needing treatment. For other populations, the DEP estimate may differentiate better between the less serious and more serious drug users clearly in need of treatment.

Small area estimation for the National Household Survey of [sic] Drug Abuse

CITATION: Folsom, R. E., & Liu, J. (1994). Small area estimation for the National Household Survey of [sic] Drug Abuse. In *Proceedings of the 1994 Joint Statistical Meetings, American Statistical Association, Survey Research Methods Section, Toronto, Ontario, Canada* (pp. 565-570). Alexandria, VA: American Statistical Association.

PURPOSE/OVERVIEW: In this paper, the authors summarized their planned approach for producing small area estimates of drug and alcohol use for selected U.S. States and metropolitan statistical areas (MSAs). The small area statistics of primary interest were population prevalence estimates of illicit drug and alcohol dependency as ascertained from responses to National Household Survey on Drug Abuse (NHSDA) questionnaire items. The authors produced separate rates for 32 demographic subpopulations defined by the three-way cross-classification of gender with four age intervals and four race/ethnicity groups where the three non-Hispanic categories excluded Hispanics. Although the authors expected most of the small area statistics for the "other races" category to be suppressed due to excessive mean squared error (MSE) estimates, the inclusion of this other races category permitted them to report statistics for Hispanics, whites, and blacks employing the same race/ethnicity definitions typically used in government data sources.

METHODS: The authors' small area estimation (SAE) strategy began by predicting block group–level drug or alcohol dependency/use rates for all the 1990 block groups in a State or MSA small area by demographic domain (the 32 gender by age by race/ethnicity subpopulations). The authors used a logistic regression predictor for the drug or alcohol dependency/use rate of domain-*d* in block group-*k* for MSA/county-*j* in State-*i*.

RESULTS/CONCLUSIONS: To quantify the uncertainty associated with the authors' nested random effect logistic model estimators, they linearized the logit-transformed versions of the statistics. For States where the fraction of the domain-*d* population that resides in sample counties was negligible, the linearized form of the equation was made proportional to a formula. To approximate another formula, the authors employed the posterior covariance matrices specified above with $\hat{D}$ replacing D. To account for the significant additional variation that may result, the authors derived a special case of Prashad and Rao's (1991) result for their three-level nested random effects model.

Prevalence of drug use in the DC metropolitan area household and nonhousehold populations: 1991

CITATION: National Institute on Drug Abuse. (1994). *Prevalence of drug use in the DC metropolitan area household and nonhousehold populations: 1991* (Technical Report #8, Washington, DC, Metropolitan Area Drug Study, DC*MADS). Rockville, MD: National Institute on Drug Abuse, Division of Epidemiology and Prevention Research. [Prepared by Bray, R. M., Kroutil, L. A., Wheeless, S. C., Marsden, M. E., & Packer, L. E.]

PURPOSE/OVERVIEW: This study from the early 1990s examined the prevalence of illicit drug, alcohol, and tobacco use among members of household and nonhousehold populations and

a combined aggregate population aged 12 or older in the District of Columbia metropolitan statistical area (DC MSA). In addition, selected characteristics of three drug-abusing subgroups in the household and aggregate populations were examined: crack cocaine users, heroin users, and needle users. The study had three methodological objectives: (1) to investigate the effect that combining data from household and nonhousehold populations had on estimates of the prevalence of drug use and number of users, (2) to determine whether the addition of nonhousehold populations allowed more detailed demographic analyses to be conducted for specific drug-using behaviors, and (3) to identify important methodological issues when combining and analyzing data from household and nonhousehold populations.

METHODS: Household population data were collected as part of the DC MSA oversample of the 1991 National Household Survey on Drug Abuse (NHSDA) (Inter-university Consortium for Political and Social Research [ICPSR] 6128). Nonhousehold population data were subsetted from the 1991 DC*MADS Institutionalized Population Study and the 1991 DC*MADS: Homeless and Transient Population Study (ICPSR 2346).

RESULTS/CONCLUSIONS: Household survey topics included age at first use, as well as lifetime, annual, and past month usage for the following drug classes: marijuana and hashish, cocaine (and crack), hallucinogens, heroin, inhalants, alcohol, tobacco, anabolic steroids, nonmedical use of prescription drugs (including psychotherapeutics), and polysubstance use. Respondents also were asked about substance abuse treatment history, problems resulting from use of drugs, perceptions of the risks involved, personal and family income sources and amounts, need for treatment for drug or alcohol use, mental health and access to care, illegal activities and arrests, and needle-sharing. Demographic data for the household population included gender, race, age, ethnicity, marital status, motor vehicle use, educational level, job status, income level, veteran status, and past and current household composition. Topics from the Institutionalized Study and the Homeless and Transient Population Study included history of homelessness, alcohol and drug treatment or counseling, illegal activities and arrests, physical health, pregnancy, mental health, mental health treatment, employment, income and expenditures, living arrangements and population movement, and specific and general drug use. Drugs covered included tobacco, alcohol, marijuana and hashish, inhalants, cocaine and crack, hallucinogens, heroin, stimulants, and tranquilizers. Data also provided information on age at first use, route of administration, polysubstance use, perceived risks, and insurance coverage. Demographic data for the nonhousehold population included gender, age, marital status, race, education, military service, and number of children and other dependents.

Breach of privacy in surveys on adolescent drug use: A methodological inquiry

CITATION: Schütz, C. G., Chilcoat, H. D., & Anthony, J. C. (1994). Breach of privacy in surveys on adolescent drug use: A methodological inquiry. *International Journal of Methods in Psychiatric Research, 4*(3), 183-188.

PURPOSE/OVERVIEW: Drug use is one of the more sensitive topics of surveys in the United States. Although there is much effort to ensure the confidentiality of surveys, there are some occasions where this confidentiality may be breached. For example, when interviews are done in

the home, a third party, such as a parent, may be present throughout the interview. The authors investigated whether this breach of privacy during National Household Survey on Drug Abuse (NHSDA) interviews affected the reporting of drug use among adolescents who participated in the survey.

METHODS: The authors used NHSDA data to test the association between the level of privacy and the reported prevalence of drug use. They selected survey participants aged 12 to 17 years old (n = 2,152) for the analysis. Interviewers provided ratings of privacy after the interview.

RESULTS/CONCLUSIONS: It was found that adolescent respondents were less likely to report their involvement with tobacco, alcohol, and illegal substances, such as marijuana and cocaine, when their parents were present during the interview than when complete privacy was reported. However, when there was a nonparent present during the interview, adolescent respondents were more likely to report their use of alcohol, tobacco, or illicit drugs than when there was complete privacy. However, in additional analyses that controlled for other variables, it was found that the presence of a person other than a parent did not have statistically significant effects on drug use reporting.

Ratio estimation of hard-core drug use

CITATION: Wright, D., Gfroerer, J., & Epstein, J. (1994). Ratio estimation of hard-core drug use. In *Proceedings of the 1994 Joint Statistical Meetings, American Statistical Association, Survey Research Methods Section, Toronto, Ontario, Canada* (pp. 541-546). Alexandria, VA: American Statistical Association.

PURPOSE/OVERVIEW: In the early 1990s, the need for accurate estimates of the size of the so-called "hard-core" drug-using population was substantial. Regardless of how it was specifically defined, this population of heavy drug users was likely to need significant resources for treatment of their drug problem and associated medical and other problems. Hard-core drug users also had been shown to be responsible for a disproportionate amount of crime.

This paper describes a method for estimating the prevalence of hard-core drug use based on the National Household Survey on Drug Abuse (NHSDA) in conjunction with outside sources and the methodology of ratio estimation. In ratio estimation, one can often obtain a better estimate of a population total if there is a known population total for a related variable. Then the estimate of the total is $X' = (x/y)*Y$, where x is the variable of interest, y is the related variable, and Y is the known population total for the related variable.

Another way of describing this method is to say that it "inflates" (i.e., gives more weight to) the drug prevalence data from the NHSDA for populations with characteristics that are known to be related to hard-core drug use but also are underestimated. In this case, it is known that the NHSDA undercounted arrestees and drug treatment populations, so the authors "ratio adjust" the NHSDA hard-core drug use estimates upward to externally derive counts of arrestees and treatment clients that are believed to be accurate.

In survey sampling theory, ratio estimation often is associated with the desire to improve the precision of an estimate. The ratio estimate will be better, in the sense that it will have a smaller variance when certain conditions are met. (See Section 6 of this 1994 paper on precision of the estimates.) However, in this application, the authors were less interested in variance reduction and more interested in bias reduction. Ratio estimates have been used for a number of years to adjust for nonresponse and to adjust to known population counts, often based on a census. This application represents an extension of those earlier uses to one in which a known population count was used to adjust NHSDA sample estimates for underreporting and undercoverage.

METHODS: The information that the authors made use of is the count of the number of persons in treatment centers for drug abuse during the past year (1992) from the National Drug and Alcoholism Treatment Unit Survey (NDATUS) (Substance Abuse and Mental Health Services Administration [SAMHSA], 1993b) and the known count of the number of arrests (for any crime other than minor traffic violations) during the past year (1991) from the Federal Bureau of Investigation's Uniform Crime Reports (UCRs) (Maguire, Pastore, & Flanagan, 1993).

RESULTS/CONCLUSIONS: A complete evaluation and comparison of the ratio estimation procedure with other methods of estimating hard-core drug use was beyond the scope of this paper. However, the authors offered some overall statements about ratio estimation: (1) Ratio estimation does not fully account for underreporting and undercoverage in the NHSDA. In particular, for the population not arrested and not in treatment, the method does not adjust for underreporting at all. Thus, the authors considered these estimates of hard-core drug use to be improvements on the generally published NHSDA estimates (using the simple expansion estimator) but still conservative estimates. (2) The ratio estimation model, as applied in this case, relied primarily on regularly updated and consistently collected data from the NHSDA, NDATUS, and UCR, and a relatively small number of easily understood assumptions. Thus, it was likely to be able to provide more reliable trend information (given constant levels of underreporting) than the previously used methods, which relied more heavily on assumptions that could change over time. (3) Because the model relied primarily on the NHSDA sample design and weighting, it was possible to develop estimates of the variances of ratio-adjusted estimates. This was generally not possible in the methods previously used.

The authors indicated that there were three primary areas for further investigation: (1) the population counts, (2) the assumptions made about the ratios used, and (3) a search for "unbiased" methods to estimate the ratio.

1995

Estimation of drug use incidence using NHSDA

CITATION: Johnson, R., Gerstein, D., & Gfroerer, J. (1995). Estimation of drug use incidence using NHSDA. In *Proceedings of the 1995 Joint Statistical Meetings, American Statistical Association, Survey Research Methods Section, Orlando, FL* (pp. 437-442). Alexandria, VA: American Statistical Association.

PURPOSE/OVERVIEW: This paper reviews the methods of data collection and statistical estimation used to estimate rates of drug use incidence in the National Household Survey on Drug Abuse (NHSDA), the primary source of data on the prevalence of drug use in the United States, and assesses possible biases in these estimated rates. NHSDA is conducted annually by the Substance Abuse and Mental Health Services Administration (SAMHSA) and is a continuing, cross-sectional, personal-interview survey of persons aged 12 or older in the civilian, noninstitutionalized population of the United States.

METHODS: This paper uses data from the 1988, 1990, 1991, 1992, and 1993 NHSDAs. There were approximately 8,000 respondents in each of the 1988 and 1990 NHSDAs and approximately 30,000 in each of the 1991, 1992, and 1993 NHSDAs. The five surveys used essentially the same drug use questions, and the 1991 to 1993 surveys oversampled the same six major metropolitan areas.

RESULTS/CONCLUSIONS: Rates of drug use incidence are critical statistics in monitoring trends in drug use in the United States. A review of data collection and estimation procedures in the NHSDA suggested three sources of possible bias in such rates: (1) differential mortality, (2) memory errors, and (3) social acceptability/fear of disclosure. The young ages of most first drug users and historical data on U.S. mortality suggested differential mortality was unlikely to be a serious source of bias in rates computed for reference periods fewer than 30 years ago. Except for the cigarette questions, NHSDA used self-administered answer sheets for all drug questions and other procedures designed to reduce social acceptability bias and respondents' perceived risks of disclosure. (The 1994 NHSDA converted the tobacco section to the self-administered format.) Preliminary analyses comparing estimates for the same reference periods computed using NHSDAs conducted in different years suggested that biases due to memory errors were small for reference periods prior to the 1990s.

1996

Repeated measures estimation of measurement bias for self-reports of drug use with applications to the National Household Survey on Drug Abuse

CITATION: Biemer, P., & Witt, M. (1996). Repeated measures estimation of measurement bias for self-reports of drug use with applications to the National Household Survey on Drug Abuse. *Journal of Official Statistics, 12*(3), 275-300.

PURPOSE/OVERVIEW: This article explores a variety of methods used to assess the validity of self-reported drug use. Using repeated measures, such as re-interview methods, test-retest, record check studies, and biological test validation, the researchers estimated measurement bias on self-reported drug use. This paper describes an improved method for estimating false-positive and false-negative self-reports in repeated measures studies.

METHODS: The authors explain that traditional repeated measures methods rely on the assumption that the second measure has considerably lower error rates than the parent survey; however, that is often not the case. In this paper, a number of methods that do not require that the second measurement be more accurate than the first were evaluated. One such methodology, the Hui-Walter method, is implemented on the National Household Survey on Drug Abuse (NHSDA) where repeated measures on several variables are available. The goal was to estimate the percentages of false negatives and false positives on several key variables to assess the validity of self-reported drug use.

RESULTS/CONCLUSIONS: This article presents a general model for estimating the proportion of misclassification in self-reported drug studies. The model was applied to studies where repeated measures of the same item are available. The Hui-Walter method produced estimates for the prevalence of drug use that are adjusted for the rate of false positives and false negatives. The next step is to use the Hui-Walter method in validating self-report errors in studies that use a biological test as a validation measure.

Aggregating survey data on drug use across household, institutionalized, and homeless populations

CITATION: Bray, R. M., Wheeless, S. C., & Kroutil, L. A. (1996). Aggregating survey data on drug use across household, institutionalized, and homeless populations. In R. Warnecke (Ed.), *Sixth Conference on Health Survey Research Methods: Conference proceedings* [Breckenridge, CO] (DHHS Publication No. PHS 96-1013, pp. 105-110). Hyattsville, MD: U.S. Department of Health and Human Services, Public Health Service, Centers for Disease Control and Prevention, National Center for Health Statistics.

PURPOSE/OVERVIEW: Since 1971, the National Household Survey on Drug Abuse (NHSDA) series has provided key information about the extent of drug use among people living in households in the United States. However, some NHSDA estimates have been criticized in

recent years because the survey has been limited in its ability to adequately represent populations that are at high risk for abusing alcohol or illicit drugs, such as incarcerated or homeless people. This paper describes an effort to combine data obtained from members of households, institutionalized populations, and homeless populations aged 12 or older into an aggregate population in the District of Columbia metropolitan statistical area (DC MSA). The paper also presents findings about drug use prevalence for both household and aggregate populations.

METHODS: This research is based on data from three surveys in the DC MSA: the DC MSA oversample of the 1991 NHSDA, the DC*MADS Institutionalized Study, and the DC*MADS Homeless and Transient Population Study. Because the populations included in these studies are generally defined in terms of residence, the overlap is very small: less than 0.5 percent of the total combined population. However, adjustments were made to avoid multiple counting of the subpopulations when producing aggregate estimates. Respondents were first classified according to the number of overlapping surveys for which they could have been selected. At most, the overlap occurred only in two of the three surveys. Analysis weights were adjusted for these respondents. The adjusted weights were summed with the final analysis weights to form multiplicity-adjusted weights for the aggregate data.

RESULTS/CONCLUSIONS: To assess the trade-offs in the use of multiplicity estimates, the variance of key estimates was examined. The authors found that the aggregate dataset produced unbiased estimates of the prevalence of illicit drug, alcohol, and cigarette use among the eligible population in the DC MSA. The authors also found that combining data from household and nonhousehold populations resulted in slightly higher prevalence estimates than from household data alone, although it had little impact on estimates within demographic subgroups for any illicit drug, alcohol, and cigarette use. However, the aggregate population totals yielded more illicit drug users. Findings from this study may underscore a potential limitation in reporting overall estimates for a large population. Such general estimates can obscure high rates of drug use or related problems among subgroups that constitute only a small percentage of the overall population.

Increasing response rates and data quality in personal interview surveys without increasing costs: An application of CQI to the NHSDA

CITATION: Burke, B., & Virag, T. (1996). Increasing response rates and data quality in personal interview surveys without increasing costs: An application of CQI to the NHSDA. In *Proceedings of the 1996 Joint Statistical Meetings, American Statistical Association, Survey Research Methods Section, Chicago, IL, Vol. 2* (pp. 1026-1031). Alexandria, VA: American Statistical Association.

PURPOSE/OVERVIEW: The consensus among survey researchers in 1996 was that there had been a decline in recent years in response rates for large, national personal visit surveys. Many national surveys required increased data collection efforts to maintain acceptable response rates (i.e., rates comparable with previous rounds [Groves & Couper, 1992]). As part of Research Triangle Institute's (RTI) continuing effort to change this trend, RTI developed a program of continuous quality improvement (CQI) for its field staff. This paper presents the history, design, implementation, and results of the first year of this CQI program that was implemented for the

1995 National Household Survey on Drug Abuse (NHSDA). Many of the features of its CQI program are readily transferable to other national personal visit surveys.

METHODS: In the summer of 1994, RTI convened a task force of individuals drawn from RTI management, NHSDA project management, survey methodologists, statisticians, and programmers/analysts to investigate data collection improvement areas for the NHSDA. The task force held a series of meetings and several focus group sessions involving the NHSDA field supervisors. Based on these discussions, RTI developed a program of CQI for NHSDA field interviewing. The objectives of this program were fourfold: (1) to increase screening and interviewing response rates to their highest levels ever; (2) to reduce costs and increase data quality; (3) to provide the field supervisors (FSs) and the field interviewers (FIs) with the tools, resources, and support to achieve unprecedented gains in quality; and (4) to appropriately reward those FSs and FIs who were responsible for the improvements in proportion to their contributions to the program's objectives.

To achieve these objectives, NHSDA management staff planned a system of performance measurement, communication, and improvement that emphasized teamwork, numerical goals, and data-driven decisionmaking. An important feature of this approach was a program of recognition and reward commensurate with team performance on the key, critical success factors that define field interviewing quality. This was implemented through a team-building concept called "Together Everyone Achieves More" (TEAM), which incorporated basic Total Quality Management and CQI concepts.

RESULTS/CONCLUSIONS: The authors saw some encouraging results at the end of the first year of the TEAM program. They had seen the best response rates since 1992 as well as a reduction in field costs per completed interview from the levels experienced during the 1994 NHSDA. They determined that continued evaluation of performance on the remainder of the 1995 NHSDA and throughout the 1996 NHSDA would allow them to evaluate whether the TEAM program would continue to show increases in response rates along with decreases in operating costs. The authors concluded that more research was required to determine whether this approach to managing national personal interview surveys could be applied to other studies outside the NHSDA project.

Social environmental impacts on survey cooperation

CITATION: Couper, M. P., & Groves, R. M. (1996). Social environmental impacts on survey cooperation. *Quality and Quantity, 30*(2), 173-188.

PURPOSE/OVERVIEW: This paper explores the effects on survey cooperation rates due to the social context and environment of the respondent household, such as urbanicity, population density, crime, and social disorganization.

METHODS: Using nonresponse data from six surveys, including the National Household Survey on Drug Abuse (NHSDA), and comparing them with census data, social context, and environmental factors were analyzed among 11,600 cases. Multivariate modeling was performed using logistic regression to analyze the difference between responders and nonresponders.

RESULTS/CONCLUSIONS: The results confirm the authors' hypotheses that highly urban areas, such as densely populated metropolitan areas, and high crime rates were associated with higher refusal rates. Controlling for household-level variables, such as household structure, race, age, number of children, and socioeconomic status, strengthened the effect of social environmental factors, but it did not eliminate them. A large percentage of the influence of environmental factors on response rates can be attributed to household-level variables. These findings show that survey participation is associated with both person- and household-level factors.

Special populations, sensitive issues, and the use of computer-assisted interviewing in surveys

CITATION: Gfroerer, J. (1996, April). Special populations, sensitive issues, and the use of computer-assisted interviewing in surveys. In R. Warnecke (Ed.), *Sixth Conference on Health Survey Research Methods: Conference proceedings* [Breckenridge, CO] (DHHS Publication No. PHS 96-1013, pp. 177-180). Hyattsville, MD: U.S. Department of Health and Human Services, Public Health Service, Centers for Disease Control and Prevention, National Center for Health Statistics.

PURPOSE/OVERVIEW: This discussion paper summarizes a session at the Sixth Conference on Health Survey Research Methods, which included papers on three separate but related areas of interest to health survey researchers: surveys of special populations, collecting data on sensitive issues, and the use of computer-assisted interviewing (CAI). Data from the National Household Survey on Drug Abuse (NHSDA) were used in the research to address the collection of sensitive data from youths. Each of these three topics received increasing attention in the early to mid-1990s.

Surveys of special populations have become more important as health planners and policymakers required data to address the health care needs of specific population subgroups. Although sometimes these data could be obtained from ongoing broad-based surveys, often it is necessary to conduct separate surveys targeting special populations. Available ongoing surveys may not have sufficient numbers of cases in the population of interest or may not even include the population in their universe (e.g., most household surveys excluded homeless persons). Many of the same methodological issues apply whether the special population is surveyed in a limited study or as part of a larger survey with broader coverage.

Surveys of sensitive issues have become more prevalent, in part to provide critical data describing emerging health problems, such as acquired immunodeficiency syndrome (AIDS) and drug abuse. These health problems require survey researchers to collect data on sensitive topics, such as sexual behavior and illegal activities. This in turn requires new and innovative methods for ensuring the validity of the data collected in surveys.

The third major area discussed in this session is the use of CAI, which was rapidly becoming the standard for all large-scale surveys by the mid-1990s. Many studies document the data collection, processing, and quality benefits of CAI. As costs continue to decline and improved technology and software become available, many surveys have been converting to CAI,

including computer-assisted telephone interviewing (CATI), computer-assisted personal interviewing (CAPI), computer-assisted self-interviewing (CASI), and audio computer-assisted self-interviewing (ACASI).

METHODS: The emergence of each of these three major areas in health survey research occurred somewhat independently and for unrelated reasons. However, it is difficult to discuss them separately. The methodological questions that they generate often overlap, and the answers given by research in these areas are sometimes meaningful only in the context of the others. For example, some topics that are sensitive for some special populations may not be sensitive for others. An example of this is alcohol use, which is thought to be sensitive for underage youths but not for adults. Similarly, some types of CAI (e.g., CASI) may work very well for most populations but may be problematic for some special populations, requiring specialized methods. And respondents' willingness to report sensitive data may vary with different types of CAI.

RESULTS/CONCLUSIONS: During the period of rapid conversion of surveys to CAI in the 1990s, it was critical that methodological research include studies of the benefits and effects on data that CAI has in surveys of special populations and on sensitive topics. The six papers presented in this session add important new knowledge to this growing body of research, and are useful to government agencies conducting surveys of special populations and on sensitive issues and that are considering the use of CAI.

The influence of parental presence on the reporting of sensitive behaviors by youth

CITATION: Horm, J., Cynamon, M., & Thornberry, O. (1996, April). The influence of parental presence on the reporting of sensitive behaviors by youth. In R. Warnecke (Ed.), *Sixth Conference on Health Survey Research Methods: Conference proceedings* [Breckenridge, CO] (DHHS Publication No. PHS 96-1013, pp. 141-145). Hyattsville, MD: U.S. Department of Health and Human Services, Public Health Service, Centers for Disease Control and Prevention, National Center for Health Statistics.

PURPOSE/OVERVIEW: A continuing challenge facing survey researchers in the mid-1990s was administering questions on respondent-perceived sensitive behaviors in a manner that minimized response bias. Negative personal behaviors could be frequently underreported or inaccurately reported because of fear of disclosure, including inadvertent disclosure to household members. Telephone and self-administered interviews, which were often used as solutions to this problem, had serious shortcomings (Aquilino, 1994; Gfroerer, 1985; Groves, 1990; Johnson, Hougland, & Clayton, 1989; Schwarz, Strack, Hippler, & Bishop, 1991). Telephone coverage was incomplete, and respondents may not have been comfortable answering questions about sensitive behaviors when there was a concern that another household member may overhear or listen in on the conversation. Further, the absence of a social relationship between the interviewer and the respondent during telephone interviews was believed to reduce the respondent's willingness to reveal personal information. Although generally producing higher reported levels of sensitive behaviors, a self-administered document had the major shortfall of requiring the respondent to have a sophistication and skill in both reading and filling out complex questionnaires. In the mid-1990s, the recently developed computer-assisted self-

interviewing (CASI) technique, with audio, overcame some of the problems of literacy and privacy, but it was not always feasible due to budgetary constraints (Mosher, Pratt, & Duffer, 1994; O'Reilly, Hubbard, Lessler, Biemer, & Turner, 1994). This paper discusses an alternative mode that was inexpensive and offered privacy for household surveys on sensitive topics.

METHODS: In 1992, the National Center for Health Statistics (NCHS) fielded the Youth Risk Behavior Survey (YRBS) as a follow-up to the National Health Interview Survey (NHIS). Because of concerns about accurate and complete reporting for the sensitive subjects included on the 1992 YRBS, the use of innovative methods to minimize risk of disclosure was explored. To ensure the privacy of the respondents so that participation and honesty could be maximized, the interviews were administered using a portable audiocassette tape player with headphones (PACTAPH), with the questions and answer categories recorded on tape. Interviewers also recorded information on the proximity of parents during the interview. Data from the 1992 National Household Survey on Drug Abuse (NHSDA), a survey in which the interviewer read the questions and the respondent marked an answer sheet, were used for comparison purposes. NHSDA and YRBS share similarities in sampling methodologies, can both be used to produce national estimates, and contain comparable questions relevant to the analysis.

RESULTS/CONCLUSIONS: This paper provides evidence consistent with findings of previous research of a relationship between the degree of privacy and the frequency of positive responses to sensitive questions by youths. The evidence suggests that the PACTAPH interviewing technique provides a greater degree of privacy from parental disclosure than do interviewer- and/or self-administered approaches. Youths clearly are concerned about parents or other household members learning of their responses and require the most secure setting and interview mode to respond honestly to sensitive questions. The use of the PACTAPH approach appears to provide the level of privacy necessary for maximum disclosure of sensitive behaviors. The findings of this research suggest that broad promises of anonymity or confidentiality are perhaps less important to honest reporting by youths than are assurances of privacy from the immediate threat of disclosure to parents.

Although most levels of reporting on sensitive behaviors appeared to increase slightly when parents were at home but out of the room, there were few significant differences within age groups, suggesting that the use of PACTAPH allayed respondent fears of disclosure. With one exception, estimates from the YRBS for three measures (smoking cigarettes in the past 30 days, ever trying marijuana and ever drinking alcohol) were higher than those from the NHSDA for three age groups (12–13, 14–15, 16–17). In addition, differences in estimates by parental proximity were larger in the NHSDA than on the YRBS, further suggesting support for the premise that the primary concern to youths is the immediate threat of disclosure to parents rather than anonymity or confidentiality.

The development and implementation of a new data collection instrument for the 1994 national household survey on drug abuse

CITATION: Office of Applied Studies. (1996, April). *Development and implementation of a new data collection instrument for the 1994 National Household Survey on Drug Abuse* (DHHS Publication No. SMA 96-3084, prepared under Contract No. 283-93-5409 for the Substance Abuse and Mental Health Services Administration). Rockville, MD: Substance Abuse and Mental Health Services Administration. [Prepared by L. A. Kroutil, L. L. Guess, M. B. Witt, L. E. Packer, J. V. Rachal, & J. C. Gfroerer]

PURPOSE/OVERVIEW: Since 1971, the National Household Survey on Drug Abuse (NHSDA) series has provided researchers and policymakers with key information about the extent of drug use and drug-related problems in the United States. Because of the importance of the survey, considerable attention has focused on ways to refine the NHSDA instrument to ensure accuracy of measurement. This report presents summaries of the background and history of the NHSDA, results of methodological studies on the NHSDA instrument, background on the implementation of the new questionnaire, and details of how the questionnaire was changed.

METHODS: The authors describe the background and rationale for the redesign of the NHSDA instrument, document the improvements made to the NHSDA instrument and the implementation plan for the new instrument, and assess the quality of the data obtained from the old and new versions of the 1994 instrument. To assess the quality of the data, the authors compare unweighted data from older versions of the 1994 NHSDA instrument with data from the new versions in terms of item nonresponse, inconsistent answers among related items, and overall nonresponse, including incomplete interviews. The authors also investigate whether differences that occur between estimates based on the old and new versions could be due to changes in the instrumentation, changes in the editing procedures, a combination of questionnaire and editing effects, the effect of statistical imputation procedures for handling missing data, or sampling error.

RESULTS/CONCLUSIONS: There was no indication that the new questionnaire had any effect on respondents' willingness to participate in the survey. The quality of data in the new instrument may be better than in the older version in many ways. The degree of missing data was lower in the new questionnaire. Also, inconsistencies between drug use recency and related variables were improved in the new instrument. Comparison of prevalence estimates from the old and new versions of the instrument for 1994 produced comparable estimates for most drugs and within most demographic subgroups. Although most differences were not statistically significant, some differences were. The higher rates of cigarette and smokeless tobacco use based on the new instrument may be related to differences in administration mode (self-administered in the new instrument vs. interviewer-administered in the old instrument). Some differences in prevalence estimates appeared to be due, at least in part, to differences in the format and content of the instruments.

Estimates of drug use prevalence in Miami from the 1991–1994 National Household Survey on Drug Abuse: Methodological report

CITATION: Witt, M., Guess, L., Gfroerer, J., & Wright, D. (1996). *Estimates of drug use prevalence in Miami from the 1991–1994 National Household Survey on Drug Abuse: Methodological report.* Research Triangle Park, NC: Research Triangle Institute.

PURPOSE/OVERVIEW: From 1991 to 1993, the National Household Survey on Drug Abuse (NHSDA) sample was expanded to provide the ability to estimate drug use prevalence in six large metropolitan areas, including Miami. The resulting estimates for the Miami area indicated substantial decreases in illicit drug use from 1991 to 1993. These data were cited as evidence of the effectiveness of the Miami Coalition, an anti-substance abuse coalition formed in early 1990. However, anomalies seen in the data led the Substance Abuse and Mental Health Services Administration's (SAMHSA's) Office of Applied Studies (OAS) and the Research Triangle Institute (RTI), the contractor conducting the NHSDA, to investigate a number of factors that could have affected the NHSDA data during the 1991 to 1993 time period.

METHODS: The investigation involved a series of analyses of the NHSDA data, with a major emphasis on the impact of Hurricane Andrew, which hit Miami in August 1992. NHSDA data were analyzed separately for areas most affected by Hurricane Andrew and for the time periods before and after the hurricane.

RESULTS/CONCLUSIONS: The evidence suggests that the NHSDA-estimated decrease in illicit drug use was not due to any change in the design of the survey, in the estimation and editing methodology, or in the weighting procedures. There was considerable evidence to suggest that Hurricane Andrew did affect NHSDA results. Rates of drug use and perceived availability of drugs dropped after Hurricane Andrew. However, the demographic characteristics of the sample population did not seem to be affected by the hurricane. Additionally, although there were small decreases in NHSDA rates of current illicit drug use in Miami before Hurricane Andrew, these decreases generally were not statistically significant. Finally, the observed decrease in lifetime prevalence estimates made it questionable to conclude that there were real decreases in past month prevalence estimates.

1997

Repeated measures estimation of measurement bias for self-reported drug use with applications to the National Household Survey on Drug Abuse

CITATION: Biemer, P. P., & Witt, M. (1997). Repeated measures estimation of measurement bias for self-reported drug use with applications to the National Household Survey on Drug Abuse. In L. Harrison & A. Hughes (Eds.), *The validity of self-reported drug use: Improving the accuracy of survey estimates* (NIH Publication No. 97-4147, NIDA Research Monograph 167, pp. 439-476). Rockville, MD: National Institute on Drug Abuse.

PURPOSE/OVERVIEW: Direct estimates of response bias for self-reports of drug use in surveys require that essentially error-free determinations of drug use be obtained for a subsample of survey respondents. The difficulty of obtaining determinations that are accurate enough for estimating validity is well documented in the literature. Methods such as specimen (e.g., hair, urine) analysis, proxy reports, and the use of highly private and anonymous modes of interview all have to contend with error rates that may be only marginally lower than those of the parent survey. Thus, any methodology for direct validity estimation must rely to some extent on approximations and questionable assumptions.

METHODS: In this chapter, the authors considered a number of methods that rely solely on repeated measures data to assess response bias. Because the assumptions associated with these approaches do not require highly accurate second determinations, they may be more easily satisfied in practice. One such method for bias estimation for dichotomous variables that is considered in some detail provides estimates of misclassification probabilities in the initial measurement without requiring that the second measure be accurate or even better than the first. This methodology does require, however, that two subpopulations exist that have different rates of prevalence but whose probabilities of false-positive and false-negative error are the same. The applicability of these methods for self-reported drug use are described and illustrated using data from the National Household Survey on Drug Abuse (NHSDA).

RESULTS/CONCLUSIONS: In this chapter, a general model for studying misclassification in self-reported drug use was presented, and the model then was extended to the case where two measurements of the same characteristic are available for the sample of respondents. For the two-measurements case, the general model required seven parameters while only three degrees of freedom were available for estimation. Thus, some additional assumptions were required to reduce the set of unknown parameters to three or less. It was shown how the assumptions typically made for test-retest, true value, improved value, and Hui-Walter methods relate to the general model. Further, it was shown how the measures of reliability, measurement bias, estimator bias, mean squared error, false-negative probability, and false-positive probability can be defined in the context of the general model and how they may be estimated under the appropriate study designs.

Finally, the use of Hui and Walter's method for estimating misclassification error based on two erroneous reports was demonstrated. The reports may be self-reports, biological tests,

administrative record values, or any other measure. For the general case of two measurements, the Hui-Walter method used maximum likelihood estimation to obtain estimates of the false-negative and false-positive probabilities associated with each measurement, as well as the error-adjusted estimates of prevalence based on both measurements. The method required that the population be divided into two domains that have markedly different prevalence estimates and that satisfy the assumption of homogeneity of error probabilities.

Substance abuse in states and metropolitan areas: Model based estimates from the 1991–1993 National Household Surveys on Drug Abuse: Methodology report

CITATION: Folsom, R. E., & Judkins, D. R. (1997). *Substance abuse in states and metropolitan areas: Model based estimates from the 1991–1993 National Household Surveys on Drug Abuse: Methodology report* (DHHS Publication No. SMA 97-3140, Methodology Series M-1). Rockville, MD: Substance Abuse and Mental Health Services Administration, Office of Applied Studies.

PURPOSE/OVERVIEW: This report presents estimates of substance abuse for 26 States and 25 metropolitan statistical areas (MSAs). These estimates were developed from data collected in the National Household Survey on Drug Abuse (NHSDA) combined with local area indicators from a variety of sources. They were produced by the Substance Abuse and Mental Health Services Administration (SAMHSA) to provide State and local area policymakers with information on the prevalence of substance abuse behaviors and problems in their local areas. These estimates were an inexpensive alternative to the direct survey approach for describing substance abuse in State and local areas. They were based on a consistent methodology across areas and were constructed so that they summed to national estimates produced by the NHSDA.

METHODS: These State and MSA estimates were the result of a comprehensive small area estimation (SAE) project that included the development of an innovative methodology based on the methods used by other Federal agencies to meet needs for small area data. The methodology employed logistic regression models that combine NHSDA data with local area indicators, such as drug-related arrests, alcohol-related death rates, and block group–level characteristics from the 1990 census that were found to be associated with substance abuse. Work also was carried out to evaluate the model used to produce these estimates.

RESULTS/CONCLUSIONS: Considering a variety of evidence, it was concluded that: (1) The SAE model produced estimates of all measures that were much better than States could achieve by simply applying NHSDA national prevalence rates for demographic subgroups to the population distribution in their States. (2) A preponderance of evidence indicated that the estimates for alcohol, cigarette, and any illicit drug use were good in that they adequately reflected both levels of use and differences across States and MSAs in the level of use. (3) For past month use of any illicit drug but marijuana, past month use of cocaine, past year dependency on illicit drugs and dependency on alcohol, and need for treatment, the limited evidence indicated that these estimates also were reasonably good. (4) For arrests, past year treatment for illicit drug use, and past year treatment for alcohol, the quality assessments resulted in mixed

findings, and these small area data may not have been good indicators of either differences between States or between MSAs, except in broad terms, or of the actual levels.

Studies of nonresponse and measurement error in the National Household Survey on Drug Abuse

CITATION: Gfroerer, J., Lessler, J., & Parsley, T. (1997). Studies of nonresponse and measurement error in the National Household Survey on Drug Abuse. In L. Harrison & A. Hughes (Eds.), *The validity of self-reported drug use: Improving the accuracy of survey estimates* (NIH Publication No. 97-4147, NIDA Research Monograph 167, pp. 273-295). Rockville, MD: National Institute on Drug Abuse.

PURPOSE/OVERVIEW: A summary of the results of a series of studies of nonresponse and measurement error in the National Household Survey on Drug Abuse (NHSDA) is given in this chapter. Two studies not previously reported, the Skip Pattern Experiment and the Census Match Study, are the primary focus of the chapter. The Skip Pattern Experiment involved a test of a modified NHSDA questionnaire that made extensive use of skip patterns in drug use questions. Compared with the standard NHSDA method, which avoided skip patterns, the modified questionnaire tended to produce lower rates of reported drug use. The Census Match Study involved linking 1990 NHSDA nonrespondent cases with data from the 1990 decennial census. Household and individual data for NHSDA nonrespondents were obtained from the census and used to characterize NHSDA nonresponse patterns in detail.

METHODS: A multilevel logistic model of response propensity identified the important predictors of nonresponse, including characteristics of the sampled person, the selected household, the neighborhood, and the interviewer.

RESULTS/CONCLUSIONS: Drug abuse surveys are particularly vulnerable to nonresponse and measurement error because of the difficulties in accessing heavy drug users and the likelihood that the illegal and stigmatized nature of drug abuse may lead to underreporting. The Skip Pattern Experiment confirmed once again that respondents' reporting of their drug use behavior was highly sensitive to the conditions under which they report. This conclusion made clear the need to proceed with great caution in interpreting differences in drug use rates obtained in different surveys. It also suggested caution in the implementation of new technologies, such as computer-assisted data collection, that undoubtedly will have some as yet unknown effect on respondents' willingness to report their drug use. The Skip Pattern Experiment may have implications in the introduction of these new technologies because one of the advantages of computer-assisted interviewing is the ease with which skips can be implemented. The Census Match Study demonstrated that response rates were not constant across various interviewer, respondent, household, and neighborhood characteristics. To the extent that rates of drug use varied by these same characteristics, bias due to nonresponse may be a problem. However, it was not always the case that low response rates occurred in conjunction with high drug use prevalence. Some populations with low response rates (e.g., older adults and high-income populations) tended to have low rates of drug use. On the other hand, some populations (e.g., residents of large metropolitan areas and men) had low response rates and high drug use rates. In

estimating overall prevalence, many of these potential sources of bias would be in opposite directions and would therefore tend to cancel each other.

Prevalence of youth substance use: The impact of methodological differences between two national surveys

CITATION: Gfroerer, J., Wright, D., & Kopstein, A. (1997). Prevalence of youth substance use: The impact of methodological differences between two national surveys. *Drug and Alcohol Dependence, 47*(1), 19-30.

PURPOSE/OVERVIEW: The purpose of this paper is to analyze the differences in the methodology used in two major federally sponsored surveys of drug use among youths.

METHODS: This study compares the prevalence rates yielded by the Monitoring the Future (MTF) study and the National Household Survey on Drug Abuse (NHSDA) by comparing the differences in response rates, precision, coverage, and data collection methods.

RESULTS/CONCLUSIONS: Although the MTF has a larger sample size of youths, an analysis of precision estimates reveals that its precision is similar to that of the NHSDA. After controlling for ages used, time of data collection, and dropouts, estimates of alcohol and drug use were significantly lower in the NHSDA than in the MTF. Although the exact cause of these differences could not be determined, the most likely reasons were the different interview environments (home vs. school), the questionnaires, and nonresponse bias in the MTF.

The validity of self-reported drug use data: The accuracy of responses on confidential self-administered answer sheets

CITATION: Harrell, A. V. (1997). The validity of self-reported drug use data: The accuracy of responses on confidential self-administered answer sheets. In L. Harrison & A. Hughes (Eds.), *The validity of self-reported drug use: Improving the accuracy of survey estimates* (NIH Publication No. 97-4147, NIDA Research Monograph 167, pp. 37-58). Rockville, MD: National Institute on Drug Abuse.

PURPOSE/OVERVIEW: Official records offer a relatively inexpensive, nonintrusive strategy for checking on the accuracy of self-reported drug use. The study reported here was designed as a criterion validity test of the National Household Survey on Drug Abuse (NHSDA) procedures. In criterion validity studies, two different measures of the same trait or experience are available: a candidate measure and an external, independent criterion measure that is treated as an error-free measure of the construct.

METHODS: In this study, the underreporting of illegal drug use was investigated in a sample of 67 former drug treatment clients by comparing their survey responses with clinic records on drug problems at time of admission. The criterion measures were based on self-reported marijuana, cocaine, hallucinogens, and heroin use. Drug treatment records were obtained from the files of publicly funded drug treatment programs in three States. The study followed the NHSDA

interviewing and questionnaire procedures closely. To avoid bias from interviewer expectations and to protect the respondent's privacy, the sample of treatment clients was embedded in a larger sample of respondents. Interviewers were not told that the respondents had been treated for drug abuse. Special sample selection directions, tailored to match the target respondent's age and gender, were used to select the former drug treatment client within the household, simulating the random selection screening instrument used in NHSDA. This analysis compared reports of past year and lifetime use of marijuana, cocaine, hallucinogens, and heroin by the former drug treatment clients with the drugs listed as problematic at time of admission to treatment. This analysis also examined factors that might have influenced the respondents' willingness or ability to respond accurately, such as the level of privacy during the interview and the amount of time between admission to the program and the interview.

RESULTS/CONCLUSIONS: The accuracy of reports compared with clinic records varied by drug, with the percentage of known users reporting their use highest for marijuana, followed by cocaine and hallucinogens, and lowest for heroin. Almost half of this sample of former treatment clients denied ever receiving drug treatment.

Recall decay and telescoping in self-reports of alcohol and marijuana use: Results from the National Household Survey on Drug Abuse (NHSDA)

CITATION: Johnson, R. A., Gerstein, D. R., & Rasinski, K. A. (1997). Recall decay and telescoping in self-reports of alcohol and marijuana use: Results from the National Household Survey on Drug Abuse (NHSDA). In *Proceedings of the 1997 Joint Statistical Meetings, 52nd annual conference of the American Association for Public Opinion Research, Norfolk, VA* (pp. 964-969). Alexandria, VA: American Statistical Association.

PURPOSE/OVERVIEW: To a large extent in the mid-1990s, the knowledge of life-cycle patterns of drug use in the United States was based on retrospective self-reports of survey respondents. Although most evidence for validity came from studies comparing self-reports of narcotics addicts with hospital and criminal justice records and with the results of urine tests (Nurco, 1985), such measures were too expensive to obtain in general population surveys. Instead, many general population surveys used the reinterview design to evaluate response error.

METHODS: The authors used a repeated cross-sectional design to evaluate response errors and analyzed changes in the distribution of responses of the same birth cohorts as measured in cross-sectional surveys conducted in different years. There were advantages and disadvantages to this method. The advantage is that pooling data from 10 National Household Surveys on Drug Abuse (NHSDAs) conducted between 1979 and 1995 allowed birth cohorts for a 16-year period at minimal expense; some disadvantage were the target population and methodological differences.

RESULTS/CONCLUSIONS: The authors used a cross-sectional design to show that NHSDA estimates of alcohol and marijuana incidence were biased downward by response error, especially at ages 10 to 14. The downward bias increased with the retention interval, making stable trend lines look increasing when estimated using a single survey. In addition, there was evidence that suggests forward telescoping, as well as some intentional concealment among older respondents.

Is "up" right? The National Household Survey on Drug Abuse [review]

CITATION: Miller, P. V. (1997). Is "up" right? The National Household Survey on Drug Abuse [review]. *Public Opinion Quarterly, 61*(4), 627-641.

PURPOSE/OVERVIEW: For almost a decade, starting in the late 1980s, politicians quoted statistics from the National Household Survey on Drug Abuse (NHSDA) to claim that drug use among teenagers was rising. However, results from the 1996 NHSDA actually showed a decline in teenage drug use. This 1997 review discusses the NHSDA methodology to analyze the assumption that drug use was rising despite statistics to the contrary.

METHODS: A number of methodological factors, such as sampling, response rates, and data collection issues, are examined to assess their impact on prevalence estimates. The NHSDA sampling frame does not include individuals who are homeless (not living in shelters), nor does it include institutionalized persons. There is no evidence on whether nonresponders to the NHSDA differ from responders in level of drug use. The NHSDA went through a number of methodological changes over the years aimed to improve survey quality and encourage reporting. These factors influence the likelihood of whether the NHSDA underreports drug use.

RESULTS/CONCLUSIONS: This examination of the NHSDA methodology suggests that the survey might have underestimated drug use. However, because no studies on the likelihood of overreporting had been conducted, it was impossible to say with conviction whether the estimates were an underreport. Validation studies do not answer all the questions surrounding data quality issues in the NHSDA; however, they do help to understand how different methods have an impact on survey estimates.

The drug abuse treatment gap: Recent estimates

CITATION: Woodward, A., Epstein, J., Gfroerer, J., Melnick, D., Thoreson, R., & Willson, D. (1997). The drug abuse treatment gap: Recent estimates. *Health Care Financing Review, 18*(3), 5-17.

PURPOSE/OVERVIEW: In the mid-1990s, it was acknowledged that the National Household Survey on Drug Abuse (NHSDA) underestimated the prevalence of drug use in the United States due to coverage error and social desirability bias. In this article, the authors attempted to correct for the underreporting by applying new estimation procedures.

METHODS: To control for sampling errors and underreporting due to social desirability in the NHSDA, three smaller data sources were used to assist with the ratio estimation. The three data sources were (1) the National Drug and Alcoholism Treatment Unit Survey (NDATUS), later known as the Uniform Facility Data Set (UFDS); (2) the Drug Services Research Survey (DSRS); and (3) the Uniform Crime Report (UCR). Ratio estimation relies on the assumption that estimates can be improved when a known value exists for a related variable. To assess the effects on the need for treatment for drug abuse, the term "need" had to be redefined.

RESULTS/CONCLUSIONS: The new definition for "need" included both a clinical and epidemiological perspective, making the definition more comprehensive and reliable for use in reporting estimates. The ratio estimation procedures improved the NHSDA estimates for the number of people in need of treatment because it corrected for coverage error and underreporting.

Ratio estimation of hardcore drug use

CITATION: Wright, D., Gfroerer, J., & Epstein, J. (1997). Ratio estimation of hardcore drug use. *Journal of Official Statistics, 13*(4), 401-416.

PURPOSE/OVERVIEW: In the mid-1990s, the prevalence of hard-core drug use was consistently underestimated in national surveys. The goal of this paper was to improve the estimates for hard-core drug use from the National Household Survey on Drug Abuse (NHSDA) by using additional sources and applying ratio estimation.

METHODS: Ratio estimation improves an estimate by comparing it with known totals for a related variable. To estimate hard-core drug use, the authors made use of two known populations related to drug use. First was the number of people in treatment centers for drug abuse, which was taken from the National Drug and Alcoholism Treatment Unit Survey (NDATUS). The second known population was the number of arrests during the past year taken from the Federal Bureau of Investigation's Uniform Crime Reports (UCRs). To get estimates of hard-core drug users, the NHSDA ratio of drug users to people in treatment was multiplied by the known total population of people in treatment. The same procedure was applied to the number of arrests during the past year. These methods yielded two estimates for the number of hard-core drug users. Another step of ratio estimation was to use the two known populations together to create one single, more accurate estimate.

RESULTS/CONCLUSIONS: The standard errors for the ratio estimation were similar to the standard errors for the simple expansion estimator that was alternatively used. The ratio estimates were considered an improvement over other estimation techniques. However, they still did not account for underreporting in people who were not arrested or in treatment centers. An advantage to the estimates provided by ratio estimation is that it could provide more than just the "bottom line" and could yield estimates for such subgroups as race, region, gender, and income.

The use of external data sources and ratio estimation to improve estimates of hardcore drug use from the NHSDA

CITATION: Wright, D., Gfroerer, J., & Epstein, J. (1997). The use of external data sources and ratio estimation to improve estimates of hardcore drug use from the NHSDA. In L. Harrison & A. Hughes (Eds.), *The validity of self-reported drug use: Improving the accuracy of survey estimates* (NIH Publication No. 97-4147, NIDA Research Monograph 167, pp. 477-497). Rockville, MD: National Institute on Drug Abuse.

PURPOSE/OVERVIEW: In the mid-1990s, levels of hard-core drug use were especially difficult to estimate because of the relative rarity of the behavior, the difficulty of locating hard-core drug users, and the tendency to underreport stigmatized behavior. This chapter presents a new application of ratio estimation, combining sample data from the National Household Survey on Drug Abuse (NHSDA) together with population counts of the number of persons arrested in the past year from the Uniform Crime Report (UCR) and the number of persons in drug treatment programs in the past year from the National Drug and Alcoholism Treatment Unit Survey (NDATUS). The population counts served as a benchmark accounting for undercoverage and underreporting of hard-core drug users.

METHODS: In this discussion, the focus was on the ratio estimate's ability to reduce bias (in particular, the undercounting of hard-core drug users in the NHSDA) given a true population value of a related variable. To make the discussion more concrete, the estimation procedure was applied to four separate, but overlapping, measures of hard-core drug use for 1992: the number of past year users of heroin, weekly users of cocaine in the past year, past year users who were more dependent on some illicit drug, and past year intravenous drug users.

RESULTS/CONCLUSIONS: Overall statements about ratio estimation for the NHSDA include the following: (1) ratio estimation did not fully account for underreporting in the NHSDA; (2) because ratio estimation could be looked at as an adjustment to the NHSDA analytic weights (which were based on a probability sample design), it provided analytic capabilities that were not possible in other methods; (3) the ratio estimation model, as applied in this case, relied primarily on regularly updated and consistently collected data from the NHSDA, NDATUS, and UCR, and a relatively small number of easily understood assumptions; and (4) because ratio estimation relied primarily on the NHSDA sample design and weighting, it was possible to develop estimates of the variances of ratio-adjusted estimates.

1998

Major design changes in the National Household Survey on Drug Abuse

CITATION: Barker, P., Gfroerer, J., Caspar, R., & Lessler, J. (1998). Major design changes in the National Household Survey on Drug Abuse. In *Proceedings of the 1998 Joint Statistical Meetings, American Statistical Association, Survey Research Methods Section, Dallas, TX* (pp. 732-737). Alexandria, VA: American Statistical Association.

PURPOSE/OVERVIEW: This paper reports on interim results of methodological research carried out to test design changes eventually introduced in the 1999 National Household Survey on Drug Abuse (NHSDA). The main focus is on a feasibility test conducted in 1996 that compared administering the survey through paper-and-pencil interviewing (PAPI) and computer-assisted interviewing (CAI). The paper also briefly describes laboratory testing of 12-month frequency of use questions and plans to increase the NHSDA sample sizes in 1999 to produce State-level estimates.

METHODS: In the 1996 CAI Feasibility Experiment, 400 interviews were carried out using either PAPI or CAI. For the CAI version, core items on substance use were administered using audio computer-assisted self-interviewing (ACASI). The CAI version consisted of two versions, one that incorporated skip patterns into the survey instrument and another that did not and thus, "mirrored" the PAPI version. Interviewer observation questions also were asked.

RESULTS/CONCLUSIONS: The main result of the 1996 CAI Feasibility Experiment was that a CAI approach to gathering NHSDA data was quite feasible and had a number of benefits to offer over the PAPI approach, including increased reporting of past year and past month marijuana and cocaine use and increased perceptions of privacy. It also was found that the skip pattern version of CAI was shorter than the CAI version that "mirrored" PAPI by about 10 minutes. The cognitive laboratory testing of the 12-month frequency of use items revealed that these questions were difficult to answer due to a long recall period and that the response categories combined total number of days with periodicity estimates. An alternative response format was developed that allowed the respondent to select the unit for reporting days of use.

Reports of smoking in a national survey: Data from screening and detailed interviews, and from self- and interviewer-administered questions

CITATION: Brittingham, A., Tourangeau, R., & Kay, W. (1998). Reports of smoking in a national survey: Data from screening and detailed interviews, and from self- and interviewer-administered questions. *Reports About Smoking: Annals of Epidemiology*, *8*(6), 393-401.

PURPOSE/OVERVIEW: Using data from the 1994 National Household Survey on Drug Abuse (NHSDA), the authors tested the primary hypothesis that using self-administered questionnaires versus interviewer-administered questionnaires would increase reports of cigarette smoking

among adolescents. The study also compared respondent and proxy responses during a brief screening interview with the responses collected during a more detailed interview later.

METHODS: Approximately 22,000 respondents aged 12 or older were interviewed using a national area probability sample. During the screening interview, basic demographic information and current smoking status were collected for all household members. Using this information, one or more respondents from the household were selected to complete the main interview. Sometimes the respondent who completed the screening interview also was selected for the main interview. Sample members were randomly assigned to receive either interviewer-administered smoking questions or self-administered smoking questions. Logistic regression was used to compare the rates of smoking by interview mode and differences in reporting of smoking between the screening and the main interview.

RESULTS/CONCLUSIONS: Although the self-administered questions showed higher reports of smoking than interviewer-administered questions, the differences were not significant for the overall population. When the analyses were restricted to adolescents, there was a marginally significant main effect for mode of interview, with self-administered questions yielding higher reporting. The proportion of current smokers reported in the screening interview was significantly lower than in the main interview. Discrepancies between the screening interview and the main interview occurred 4 times more often when the screening respondent was different from the interview respondent. This supports the notion that the screening respondent was a proxy and was not necessarily cognizant of the smoking status of the rest of the household. When the same person was interviewed in the screener and in the main interview, the highest rate of discrepancy occurred in the adolescent age group, supporting the hypothesis that smoking is sensitive for adolescents but not adults. In addition, asking multiple questions about smoking leads to higher reports of smoking, particularly in respondents who may be reluctant to admit their smoking status.

Testing ACASI procedures to reduce inconsistencies in respondent reports in the NHSDA: Results from the 1997 experimental field test

CITATION: Caspar, R. A., Lessler, J. T., & Penne, M. A. (1998). Testing ACASI procedures to reduce inconsistencies in respondent reports in the NHSDA: Results from the 1997 experimental field test. In *Proceedings of the 1998 Joint Statistical Meetings, American Statistical Association, Survey Research Methods Section, Dallas, TX* (pp. 750-755). Alexandria, VA: American Statistical Association.

PURPOSE/OVERVIEW: One of the most significant potential benefits of converting the National Household Survey on Drug Abuse (NHSDA) to a computer-assisted format is the opportunity to resolve inconsistent data at the time of the interview rather than editing the data to deal with inconsistencies after the fact. However, maintaining the privacy benefits of the audio computer-assisted self-interviewing (ACASI) component of the interview requires that the respondent be able to resolve inconsistencies for many items on his or her own. Thus, one of the goals of the NHSDA conversion work was to develop a method for resolving inconsistent data that the respondent could easily understand and complete without significant intervention by the interviewer.

METHODS: Based on the authors' own hypotheses about how inconsistencies should be identified and resolved, they developed a resolution methodology that combined two components. First, at the verify stage, respondents are asked whether an answer they have entered is in fact correct. So, for example, when a 20-year-old respondent reports that she was 51 the first time she drank alcohol (a clearly inconsistent answer), the computer was programmed to verify that this information was correct. If the respondent indicated that the information was incorrect, she was routed back to answer the question again (perhaps this time entering the age of her first drink as 15). A second component incorporates the resolution of seemingly inconsistent answers. For example, a respondent who indicated drinking alcohol on 15 days in the past 12 months, but then reported drinking alcohol on 25 days in the past 30 days, first would be asked to verify the last entry keyed. If the respondent indicated that the entry was correct, then he or she was routed to a question that identified the inconsistency and provided the respondent with an opportunity to correct one or both of the entries.

RESULTS/CONCLUSIONS: Based on the results reported, the authors feel that the inconsistency resolution methodology employed in the 1997 CAI Field Experiment was successful. The methodology improved the consistency of the data collected without adversely affecting respondent cooperation or burden. Using this methodology in future implementations of the NHSDA will allow the Substance Abuse and Mental Health Services Administration (SAMHSA) to capitalize on the numerous benefits of the ACASI technology while minimizing one of the potential pitfalls—that respondent errors and inconsistencies are not identified and corrected at the time of interview. As of 1998, work is under way to incorporate inconsistency resolution screens into the 1999 computer-assisted personal interviewing (CAPI) and ACASI NHSDA instrument. The authors anticipate that future research in this area will be conducted to determine whether respondents can resolve inconsistencies between items that are not closely spaced in the interview. If this proves to be possible, the authors feel that future NHSDA instruments may incorporate inconsistency resolution screens of this type as well.

Estimating substance abuse treatment need from a national household survey

CITATION: Epstein, J. F., & Gfroerer, J. C. (1998, May). Estimating substance abuse treatment need from a national household survey. In J. Epstein (Ed.), *Analyses of substance abuse and treatment need issues* (DHHS Publication No. SMA 98-3227, Analytic Series A-7, pp. 113-125). Rockville, MD: Substance Abuse and Mental Health Services Administration, Office of Applied Studies.

PURPOSE/OVERVIEW: For purposes of planning future demand on the health care system in general and the substance abuse treatment system in particular, it is important to be able to develop estimates of the number of people needing treatment for substance abuse (i.e., treatment need) on a regular basis. In order to do this, the measurement of treatment need must distinguish low-intensity substance use from drug use that requires intervention. The overall prevalence of substance use is a poor absolute measure of problem substance use. Estimating treatment need is a difficult problem. Drug and alcohol consumption patterns and their consequences are complicated and dynamic. The modalities and philosophies of treatment are diverse. The applicability of even well-ensconced and tested diagnostic criteria must be reestablished as new drugs and ways of administering them appear. Measuring treatment need involves both a

scientific and clinical understanding of the substance abuse problem. Estimating how many people in the general population need treatment is a different problem from diagnosing the need for treatment for an individual based on history taking, physical examination, and information in previous records. In most household surveys, it would be impractical to perform physical examinations or take a detailed history. In addition, household surveys may not fully cover key populations that may have a significant number of people needing treatment, such as incarcerated and homeless individuals. The main focus of this paper describes attempts in the 1990s to estimate treatment need based on the National Household Survey on Drug Abuse (NHSDA). The NHSDA is potentially a valuable source for estimating treatment need because it is the only ongoing large national survey of substance use covering most of the population. Thus, if a reasonable method for estimating treatment need from the NHSDA could be developed, it would provide more timely estimates than other sources and could be used to measure changes in treatment need over time.

METHODS: To estimate the effects of changes in the questionnaire and variable construction, both the new and old versions of the survey were fielded in 1994. The "old" version of the survey, which was actually the same as the 1993 questionnaire, is referred to as 1994-A, while the "new" version of the survey is referred to as 1994-B.

RESULTS/CONCLUSIONS: Comparisons between the new questionnaire (1994-B) and the old questionnaire (1994-A) showed that estimates of the rate of lifetime illicit drug use from the new questionnaire were approximately 8 percent less than estimates from the old questionnaire. Estimates of the rate of past year illicit drug use were approximately 12 percent less with the new questionnaire. On the other hand, estimates of the rate of past month illicit drug use were approximately 5 percent higher with the new questionnaire. The new questionnaire produced approximately 29 percent higher estimates of the rate of weekly cocaine use and approximately 6 percent higher estimates of the rate of daily marijuana use.

Nonresponse in household interview surveys

CITATION: Groves, R. M., & Couper, M. P. (1998). *Nonresponse in household interview surveys*. New York: Wiley.

PURPOSE/OVERVIEW: The 1990 National Household Survey on Drug Abuse (NHSDA) was one of six large Federal or federally sponsored surveys used in the compilation of a dataset that then was matched to the 1990 decennial census for analyzing the correlates of nonresponse. In addition, data from surveys of NHSDA interviewers were combined with those from these other surveys to examine the effects of interviewer characteristics on nonresponse.

METHODS: Information from the NHSDA, such as interviewer notes, segment listing sheets, and segment maps, as well as other data sources were used to match sampled housing units from the 1990 NHSDA with housing units from the 1990 decennial census. Variables on household characteristics from the 1990 census and area characteristics were used as predictors of cooperation on the NHSDA and the other surveys in the dataset. For characteristics of interviewers, a survey was administered to NHSDA interviewers during training. Among the items asked about were (1) education; (2) primary reason for their interviewing job and the

existence of another paid job; (3) experience, including years worked and number of organizations worked for in the last 5 years; and (4) questions regarding interviewer attitudes and expectations. These variables were used as predictors of cooperation on the NHSDA as well as the other surveys.

RESULTS/CONCLUSIONS: Overall, 97.2 percent of NHSDA sample housing units (a total of 4,619 units) were successfully matched to the decennial census, with 97.8 percent of interviewed housing units (n = 1,407) and 96.7 percent of nonresponse housing units (n = 2,681) successfully matched. A total of 280 NHSDA interviewers completed the survey of interviewer characteristics. The main findings were as follows: (1) Once contacted, those with lower socioeconomic status were no less likely to cooperate than those with higher socioeconomic status; there was instead a tendency for those in high-cost housing to refuse survey requests, which was partially accounted for by residence in high-density urban areas. (2) The tendency for persons in military service and in non-English-speaking households to cooperate appeared to be a function of household composition. (3) Households with children or young adults were more likely to cooperate; single-adult households were less likely to cooperate. (4) After controlling for the effects of household size, households with elderly residents tended to cooperate. (5) The tendency of racial and ethnic minorities to participate was accounted for by lower socioeconomic status. (6) Densely populated, high-crime urban areas had lower cooperation rates, which could be accounted for by other ecological variables, as well as different household compositions in urban versus rural areas. (7) Interviewers with more interviewing experience tended to achieve higher cooperation rates than those with less experience, but it was not clear if this was due to higher attrition among less successful interviewers or additional skills gained by experienced interviewers. (8) There was some evidence that interviewers with higher levels of confidence in their ability to gain participation achieved higher cooperation rates.

Adjusting survey estimates for response bias: An application to trends in alcohol and marijuana use

CITATION: Johnson, R. A., Gerstein, D. R., & Rasinski, K. A. (1998). Adjusting survey estimates for response bias: An application to trends in alcohol and marijuana use. *Public Opinion Quarterly, 62*(3), 354-377.

PURPOSE/OVERVIEW: Lifetime prevalence estimates and trends in drug use are primarily based on self-reports. Validity testing usually compares the reports given by special populations with administrative records and biochemical tests. However, these types of validity measures are not possible in a general population study; therefore, most studies rely on test-retest procedures.

METHODS: Instead of the traditional reinterview method for assessing validity, this study used a repeated cross-sectional design. This method utilized birth cohorts and examined changes in the distribution of their responses over individual cross-sectional surveys. Using data from 9 years of the National Household Survey on Drug Abuse (NHSDA) from 1982 to 1995, trends in birth cohorts were evaluated over 13 years. Changes in survey design in the NHSDA over the 13 years may have introduced bias in assessing response error.

RESULTS/CONCLUSIONS: Following birth cohorts using a repeated cross-sectional design is cost-effective in assessing self-report error in studies of drug and alcohol use. Using an exponential decay model of response bias revealed lower reporting of drug use as the time between the drug use and survey response increased. This bias distorts trends in alcohol and drug use, making stable drug use trends misleadingly appear increasing. Further analysis revealed that respondents tended to forward telescope the age for first alcohol use. In addition, intentional nondisclosure of marijuana use at a young age increased as the respondent increases in age, which led to underreporting of lifetime prevalence rates for marijuana use.

Development of computer assisted interviewing procedures for the National Household Survey on Drug Abuse (NHSDA): Design and operation of the 1997 CAI field experiment

CITATION: Lessler, J. T., Witt, M., & Caspar, R. (1998). Development of computer assisted interviewing procedures for the National Household Survey on Drug Abuse (NHSDA): Design and operation of the 1997 CAI field experiment. In *Proceedings of the 1998 Joint Statistical Meetings, American Statistical Association, Survey Research Methods Section, Dallas, TX* (pp. 738-743). Alexandria, VA: American Statistical Association.

PURPOSE/OVERVIEW: A large-scale field experiment that examined the use of computer-assisted interviewing (CAI) for the National Household Survey on Drug Abuse (NHSDA) was conducted in the last quarter of 1997. The design of the 1997 field experiment was based on the results of a 1996 CAI feasibility experiment and subsequent cognitive laboratory testing, power calculations, and discussions as to the operational feasibility of various designs. The authors compared alternative versions of the audio computer-assisted self-interviewing (ACASI) portion of the CAI interview in a factorial design.

METHODS: The authors compared these alternatives with each other and with the results from the methodology employing a combination of a paper-and-pencil interviewing (PAPI) and self-administered answer sheets. They conducted the experiment in the fourth quarter of 1997 and used the Quarter 4 1997 NHSDA survey results as a comparison group.

RESULTS/CONCLUSIONS: The screening and interview response rates in the 1997 experimental field test were lower than those achieved in the main study. The overall screening response rates was 86.8 percent, which was about 7 percent lower than that achieved in similar areas in the national NHSDA. About 2.5 percent of this shortfall was due to the failure to obtain access to restricted housing; 3.5 percent was due to increased refusals. It is unlikely that the electronic screener contributed to the failure to obtain access to restricted housing. However, the authors were not able, from this study alone, to verify that using the Newton screener did not contribute to increased refusals to the screening.

Discussion of issues facing the NHSDA

CITATION: Miller, P. V. (1998). Discussion of issues facing the NHSDA. In *Proceedings of the 1998 Joint Statistical Meetings, American Statistical Association, Survey Research Methods Section, Dallas, TX* (p. 762). Alexandria, VA: American Statistical Association.

PURPOSE/OVERVIEW: The papers in this session report on the results of a field test that compared paper-and-pencil questionnaires with several versions of a computer-assisted interviewing (CAI) measurement of drug use. The study discussed is one of many studies conducted to assess the effects of converting the National Household Survey on Drug Abuse (NHSDA) from paper-and-pencil interviewing (PAPI) to CAI.

METHODS: N/A.

RESULTS/CONCLUSIONS: The results support a priori views that CAI has the potential to improve data quality in NHSDA. Findings include indications that a streamlined CAI version of the questionnaire can produce higher reporting of drug use and that inconsistency checking built into the CAI instrument can improve reporting. The authors also noted the success of ACASI technology in providing respondent privacy, handling complex skip patterns, and resolving inconsistencies between responses.

Effects of experimental audio computer-assisted self-interviewing (ACASI) procedures on reported drug use in the NHSDA: Results from the 1997 CAI field experiment

CITATION: Penne, M. A., Lessler, J. T., Bieler, G., & Caspar, R. (1998). Effects of experimental audio computer-assisted self-interviewing (ACASI) procedures on reported drug use in the NHSDA: Results from the 1997 CAI field experiment. In *Proceedings of the 1998 Joint Statistical Meetings, American Statistical Association, Social Statistics Section, Dallas, TX* (pp. 744-749). Alexandria, VA: American Statistical Association.

PURPOSE/OVERVIEW: The 1997 computer-assisted interviewing (CAI) field experiment evaluated the impact on reported drug use of using alternative versions of an audio computer-assisted self-interviewing (ACASI) version of the National Household Survey on Drug Abuse (NHSDA).

METHODS: Alternative versions of the ACASI questionnaire were examined using a factorial design conducted during the fourth quarter of 1997. A subsample of the Quarter 4 national NHSDA, which used a combination of a paper-and-pencil interviewing (PAPI) version and self-administered questionnaire (SAQ) answer sheets, comprised the control group for the study. A fuller description of the design can be found in Lessler, Witt, and Caspar (1998) in this volume. In this paper, the authors examined reported drug use and compared the experimental ACASI factors to each other and to the control group. They also presented information on overall differences between ACASI and the control group.

RESULTS/CONCLUSIONS: The prevalence data indicate that ACASI will yield higher estimates of drug use, particularly for youths. The debriefing data indicate that this is due to the privacy-enhancing features of the method.

Drug abuse treatment: Data limitations affect the accuracy of national and state estimates of need

CITATION: U.S. General Accounting Office. (1998). *Drug abuse treatment: Data limitations affect the accuracy of national and state estimates of need* (GAO/HEHS-98-229, Report to Congressional Requesters, Health, Education, and Human Service Division). Washington, DC: U.S. Government Printing Office.

PURPOSE/OVERVIEW: The Federal Government annually provides approximately $3 billion for drug abuse prevention and treatment activities. However, determining the need for treatment services for the general population, as well as for specific subpopulations, may be problematic due to limitations in national and State data on treatment need. This report describes the Substance Abuse and Mental Health Services Administration's (SAMHSA) efforts to estimate drug abuse treatment need on a national basis and obtain State estimates of drug abuse treatment need.

METHODS: The authors interviewed and obtained documents from officials in SAMHSA's Center for Substance Abuse Treatment (CSAT), the Office of Applied Studies (OAS), and the Office of the Administrator. The authors also held discussions with experts in the substance abuse research community. They reviewed needs assessment information submitted by States. In addition, the authors attended a CSAT-sponsored workshop that included all States with current State Treatment Needs Assessment Program (STNAP) contracts in which States reported on their needs assessment studies.

RESULTS/CONCLUSIONS: Reliable assessments of treatment need are an essential component to accurately target treatment services. Although SAMHSA is improving its national estimates through the expansion of the National Household Survey on Drug Abuse (NHSDA) sample, the survey is still likely to underestimate treatment need. Also, STNAP's goals to help States develop estimates of treatment need and improve State reporting of need data have not been fully accomplished. Even though States are required to provide estimates of treatment need as part of their block grant applications, not all States report this information, and some of the data reported are inaccurate. SAMHSA recognizes the need to increase State reporting and has set a target for increasing the number of States that provide the information. It also recognizes that the overall quality of the data reported is problematic.

A comparison of computer-assisted and paper-and-pencil self-administered questionnaires in a survey on smoking, alcohol, and drug use

CITATION: Wright, D. L., Aquilino, W. S., & Supple, A. J. (1998). A comparison of computer-assisted and paper-and-pencil self-administered questionnaires in a survey on smoking, alcohol, and drug use. *Public Opinion Quarterly, 62*(3), 331-353.

PURPOSE/OVERVIEW: This article assesses the impact of using computer-assisted self-interviewing (CASI) to collect self-report data on sensitive questions about drugs, alcohol, and mental health. The study also measures the interaction that computerization and respondent characteristics have on the survey responses.

METHODS: Using a national multistage area probability sample, 3,169 interviews were conducted with respondents aged 12 to 34. Approximately two thirds completed the computer self-administered questionnaire (SAQ), and the other one third completed a paper-and-pencil interviewing (PAPI) SAQ. The interviews used the drug and mental distress questions from the 1995 National Household Survey on Drug Abuse (NHSDA) questionnaire. Ordinary least squares regression was used to measure the effect of mode on prevalence estimates.

RESULTS/CONCLUSIONS: For all variables except for tobacco use, the computer SAQ produced higher estimates than the paper SAQ for adolescents. The greatest effects were seen on the most sensitive items. This supports the notion that adolescents are more familiar with and have a better understanding of computers than adults, which makes them more comfortable using the computer SAQ. There was a significant mode by mistrust interaction in older respondents for alcohol and drug use. Respondents who were mistrustful had higher estimates of drug and alcohol use in the PAPI SAQ than in the CASI SAQ.

Hierarchical models applied to the National Household Survey on Drug Abuse (NHSDA)

CITATION: Wright, D., & Zhang, Z. (1998). Hierarchical models applied to the National Household Survey on Drug Abuse (NHSDA). In *Proceedings of the 1998 Joint Statistical Meetings, American Statistical Association, Survey Methods Section, Dallas, TX* (pp. 756-761). Alexandria, VA: American Statistical Association.

PURPOSE/OVERVIEW: A fairly comprehensive discussion of the methodology in hierarchical modeling was introduced in *Hierarchical Linear Models: Applications and Data Analysis Methods* (1992) by Anthony S. Bryk and Stephen W. Raudenbush. A significant part of that book focused on hierarchical structures in the field of education and the necessity of using an estimation methodology that reflects the structure of the data. A natural question is how that methodology might apply to the drug use field. Much of the drug use analysis research over the years has utilized simple logistic regression to estimate relationships between drug use and a variety of person-level variables. The authors wanted to explore how hierarchical models could be applied to the National Household Survey on Drug Abuse (NHSDA) data and what the consequences would be of ignoring the hierarchy. Since its inception, the NHSDA has been a multiphase stratified sample of primary sampling units (counties or groups of counties), segments (blocks or block groups), households, and individuals. The result of this complex nested design is that using traditional techniques of variance estimation that ignore the clustering of sample cases and the use of sample weights tends to overstate the significance of many findings. A number of statistical packages have been developed using Taylor series methods or replication methods, such as SUDAAN® (RTI) and WESVAR® (WESTAT), that take this structure into account. There was no widely used software for hierarchical mixed models, however, until hierarchical linear modeling (HLM).

METHODS: This paper explored some of the special circumstances in applying hierarchical models to NHSDA. The authors decided to use the data from six oversampled cities from the 1991 to 1993 NHSDAs. With a sample of approximately 2,500 persons per city for each of the 3 years, there was a total sample of about 45,000 persons. Each of the six cities (Washington, DC; Chicago; Miami; New York City; Denver; and Los Angeles) was selected with certainty, and within each city, segments were sampled at the first stage. Within segment, usually about 10 to 12 persons aged 12 or older were selected. If the authors wanted to make inference to the collection of the six large cities as typical large cities, then they could consider four levels of hierarchy: city, segment, family, and person. Two issues arose immediately: Were there sufficient number of observations at each hierarchical level to support the analysis, and was it necessary to include all four levels of hierarchy in the estimation? By the latter, it is meant, "Was there sufficient variability at each stage to be necessary for inclusion?"

RESULTS/CONCLUSIONS: In the analysis, a significant portion of the variation (about 80 percent or more) was attributed to person-level variation. The person-level variation really included the variation between families as well. Because family-level variables, such as family drug use, parental attitudes about drug use, and family management practices and family conflict, had been identified as risk factors for adolescent drug use, it was possible that a significant amount of what had been labeled as person-level variation was really between-family variation. These components of variance needed to be estimated. The issue of small sample sizes at the family level in particular needed to be explored. In the 1997 and later NHSDAs, it was attempted to collect data on a random sample of pairs of persons within a sample of households. Using Goldstein's formula for the impact of this vis-a-vis ordinary least squares (OLS), the (n-1) would be equal to 2, and the inflation factor above and beyond simple random sampling would be (1+rho). For a large rho, this might make a considerable difference. If the level-2 sample sizes were larger, then even a small rho could make a difference. So, it was important to understand that any conclusions with respect to the use of HLM versus OLS applied only to the NHSDA and only with the current NHSDA sample sizes. The impact on other datasets or on the NHSDA with different sample sizes could be quite different. Although the variables that had been analyzed were interesting, there was a need to consider other drug-related variable scales and to extend the analysis to dichotomous data, such as past year use of marijuana or any illicit drug.

1999

Usability testing for self-administered survey instruments: Conversion of the National Household Survey on Drug Abuse as a case study

CITATION: Caspar, R., & Barker, P. (1999). Usability testing for self-administered survey instruments: Conversion of the National Household Survey on Drug Abuse as a case study. In R. Banks, C. Christie, J. Currall, J. Francis, P. Harris, B. Lee, J. Martin, C. Payne, & A. Westlake (Eds.), *ASC '99: Leading survey & statistical computing into the new millennium. Proceedings of the 3rd Association for Survey Computing International Conference* (pp. 79-89). Chesham, Buckinghamsire, United Kingdom: Association for Survey Computing.

PURPOSE/OVERVIEW: Computer-assisted interviewing (CAI) has prompted many changes in the way survey instruments are designed and administered. For example, computer-assisted instruments can contain significantly more complex routings than those developed for paper administration. At the same time, however, computer-assisted instruments must be designed so that they can be easily navigated by the end user. When, as in the case of self-administered questionnaires, the end user is a survey respondent who may have little or no experience with computers and may not be particularly motivated to complete the survey task, it is especially important that the computer-assisted instrument be designed to be both easily understood and easily completed. This paper describes the usability testing that has been completed as part of the work to convert the National Household Survey on Drug Abuse (NHSDA) from a paper-and-pencil interviewing (PAPI) instrument to a CAI instrument with an extensive audio computer-assisted self-interviewing (ACASI) component.

METHODS: Forty subjects were recruited to take part in the testing. Vignettes were used to simulate response inconsistencies in order to reduce costs. Respondents first were asked to verify their last response, with responsibility for the inconsistency not explicitly placed on the respondent. Original responses were displayed on screen, and respondents were asked to identify the incorrect response. The study did not address the issue of why respondents give inconsistent responses or how they would feel to have their inconsistencies identified by the computer.

RESULTS/CONCLUSIONS: An estimated 28 percent of respondents triggered a check, 45 percent of these respondents triggered a true inconsistency, and youth respondents accounted for 60 percent of the true inconsistencies triggered. Respondents triggered an average of 1.2 true inconsistencies and 1.2 "verify only." Eighty-one percent of true inconsistencies were resolved. There was a very low rate of "Don't Know" or "Refused" responses, no significant increase in interview length, and no significant increase in breakoff rates. Respondents were slightly more likely to report that they gave accurate and complete answers. Respondents also were slightly more likely to report difficulty using the computer, and interruptions and distractions were the most common reason for inconsistent data based on interviewer reports. In conclusion, usability testing improved the NHSDA instrument fielded in 1999. The approach to developing usable ACASI instruments was not so different from developing usable self-administered PAPI questionnaires. Usability testing requires planning—the researcher must allocate additional time during the development phase.

Population coverage in the National Household Survey on Drug Abuse

CITATION: Chromy, J. R., Bowman, K. R., Crump, C. J., Packer, L. E., & Penne, M. A. (1999). Population coverage in the National Household Survey on Drug Abuse. In *Proceedings of the 1999 Joint Statistical Meetings, American Statistical Association, Survey Research Methods Section, Baltimore, MD* (pp. 576-580). Alexandria, VA: American Statistical Association.

PURPOSE/OVERVIEW: This paper presents a brief description and history of the National Household Survey on Drug Abuse (NHSDA), a description of the household screening process, definitions of the coverage terms reported in this paper, a summary of overall measures across recent years, and some results by selected domains.

METHODS: For the purposes of this paper, the authors take a broad view of survey coverage in household screening surveys. They include weighted person response rates, weighted household response rates, and a measure of frame coverage based on recent experience from the NHSDA.

RESULTS/CONCLUSIONS: NHSDAs provide a large sample base for studying overall coverage rates, including those associated with both nonresponse error and frame coverage error. Results presented on frame coverage have been based on comparisons of nonresponse-adjusted estimates to external data sources. Because both are subject to error, some of the frame coverage rates need to be interpreted with caution. In particular, the Hispanic population coverage is shown to be higher than other groups, but that may be an artifact of the intercensal projections for Hispanics rather than a true measure of frame coverage. Screening response, interview response, and frame coverage have been discussed separately as if they were independent processes. In fact, this is not the case. Poor screening response can have differential effects on frame coverage if the nonscreened dwelling units tend to be ones that have a higher proportion of a certain domain (e.g., younger persons or blacks). This confounding of the screening coverage and frame coverage occurs because the development of dwelling unit rosters within dwelling units is actually the final stage of frame development. This confounding of effects may provide further support for studying all forms of undercoverage—screening nonresponse, interview nonresponse, and frame noncoverage—as an overall phenomenon.

You want 75,000 unique maps by WHEN?

CITATION: Curry, R. (1999). You want 75,000 unique maps by WHEN? In *Proceedings of the Environmental Systems Research Institute (ESRI) User Conference, San Diego, CA, July 26-30, 1999*. Redlands, CA: Environmental Systems Research Institute (ESRI).

PURPOSE/OVERVIEW: Research Triangle Institute (RTI) has been contracted to run and manage the National Household Survey on Drug Abuse (NHSDA) for the Substance Abuse and Mental Health Services Administration (SAMHSA). The 1999 NHSDA was the largest survey RTI had managed up to that date. Because of the size of the survey and the requirement for large numbers of accurate maps in a relatively short period of time, geographic information system (GIS) technology became an important part of the household survey process. The Survey Automated Mapping System (SAMS) and Setting and Zooming (SAZ) were developed to satisfy the mapping requirements of the survey. A total of 7,200 sample segments were chosen from

every State in the country. Over 1,200 field interviewers (FIs) needed to know where to administer the survey in those 7,200 geographic locations.

METHODS: For each segment, a set of maps was generated showing county, tract, segment, and block boundaries to help survey takers collect data from the correct locations. Over a 10-week period, an automated mapping system generated over 75,000 of these unique maps for the FIs.

RESULTS/CONCLUSIONS: By using a high-powered UNIX workstation, ARC/INFO software, an online TIGEr/Line database, along with data files and driver files, one person was able to use SAMS to produce high volumes of maps in a relatively short period of time while minimizing the number of errors and reruns. Other survey staff interacted with the mapping information and assisted in creating the data files for SAMS as well as producing zoom maps when needed. The field survey staff used these maps as a means to locate the survey area and to assist in the interview process.

Validity of drug use reporting in a high-risk community sample: A comparison of cocaine and heroin survey reports with hair tests

CITATION: Fendrich, M., Johnson, T. P., Sudman, S., Wislar, J. S., & Spiehler, V. (1999). Validity of drug use reporting in a high-risk community sample: A comparison of cocaine and heroin survey reports with hair tests. *American Journal of Epidemiology, 149*(10), 955-962.

PURPOSE/OVERVIEW: Toxicological tests often performed on hair specimens are the standard for assessing the accuracy of self-reported drug use data. However, most studies of this sort use specialized populations, such as inmates or people in drug use treatment programs. The purpose of this study was to use toxicological results to assess the accuracy of self-reports in a community sample.

METHODS: Using the procedures and instrument from the 1995 National Household Survey on Drug Abuse (NHSDA), respondents were assigned randomly into one of two conditions. The first condition was the control group and replicated the NHSDA procedures almost entirely. The second condition was the experimental group, which utilized a cognitive interviewing procedure. Respondents in both groups were not offered an incentive for cooperating in the survey. However, after respondents completed the survey, they were offered a $10 incentive for a specimen of their hair. Hair samples were given by 322 respondents

RESULTS/CONCLUSIONS: The results demonstrated that the toxicological tests revealed higher prevalence estimates for drug use than the survey measures. Respondents were more likely to report heroin use than cocaine use. In addition, hard-core users who had the most concentration of cocaine in their hair were more likely to report drug use than others. Respondents also were more inclined to report lifetime use than recent use, indicating that lifetime self-report estimates may be more reliable than recent-use self-reports. The findings of this study are limited by the possibility that the $10 incentive disproportionately motivated drug users to participate.

Measuring interstate variations in drug problems

CITATION: McAuliffe, W. E., LaBrie, R., Lomuto, N., Betjemann, R., & Fournier, E. (1999). Measuring interstate variations in drug problems. *Drug and Alcohol Dependence, 53*(2), 125-145.

PURPOSE/OVERVIEW: This paper describes the process used to create the Drug Problem Index (DPI), a measure that combined a collection of drug abuse indicators. The DPI was designed to compare severity of drug abuse and dependence problems across States.

METHODS: All measures included in the index came from empirical and theoretical evidence of connections to drug abuse and incorporated information from the National Drug and Alcoholism Treatment Unit Survey (NDATUS), the National Association of State Alcohol and Drug Abuse Directors (NASADAD), the Federal Bureau of Investigation's (FBI) Uniform Crime Reporting (UCR) system, and the National Household Survey on Drug Abuse (NHSDA). Measures were combined into an index by summing their *z*-scores using nominal weights. Measures for drug abuse severity had to be tested for reliability and validity to be included in the index. The measures were tested for both construct validity and external validity. Comparisons of the DPI with other measures of drug abuse also were conducted.

RESULTS/CONCLUSIONS: Three components for the DPI were selected: drug-coded mortality, drug-defined arrest, and drug-treatment client rates. Results of the reliability analysis revealed that the index demonstrated stable estimates from year to year. The index also demonstrated convergent and construct validity. In addition, the DPI correlated with the block grant drug need index. However, it did not correlate with other measures of drug abuse from NHSDA. Neither the NDATUS drug-treatment-only client rate nor the NASADAD drug admissions rate had a significant correlation with either the NHSDA direct or model-based estimates of the past year drug treatment rate.

2000

Use of alternating logistic regression in studies of drug-use clustering

CITATION: Bobashev, G. V., & Anthony, J. C. (2000). Use of alternating logistic regression in studies of drug-use clustering. *Substance Use & Misuse, 35*(6-8), 1051-1073.

PURPOSE/OVERVIEW: This paper aims to provide readers with a detailed review of the alternating logistic regression (ALR) methodology as a tool for studying clustering and community-level drug involvement.

METHODS: The authors applied the ALR analysis method to a publicly available national dataset from the National Household Survey on Drug Abuse (NHSDA) in order to examine the factors of clustered drug involvement in the United States. The variables of interest were measures of marijuana use on a weekly or more frequent basis. The ALR method can take into account sociodemographic characteristics providing estimates for different subgroups.

RESULTS/CONCLUSIONS: Clustering of drug use was apparent using the ALR methodology. Clustering was most notable in daily drug use and slightly less notable with weekly drug use. Sociodemographic characteristics of the individual, such as age, gender, income, and race, did not affect the estimates for marijuana use. However the authors' approach was unable to find strong determinants or factors that caused the clustering. They concluded that an analysis with a dataset that provided more indicators on the economic and social conditions of the neighborhood would be useful in further understanding the causes of clustering in drug use.

Experience with the generalized exponential model (GEM) for weight calibration for NHSDA

CITATION: Chen, P., Penne, M. A., & Singh, A. C. (2000). Experience with the generalized exponential model (GEM) for weight calibration for NHSDA. In *Proceedings of the 2000 Joint Statistical Meetings, American Statistical Association, Survey Research Methods Section, Indianapolis, IN* (pp. 604-607). Alexandria, VA: American Statistical Association.

PURPOSE/OVERVIEW: The National Household Survey on Drug Abuse (NHSDA) is designed to estimate prevalence of both licit and illicit drug use in the United States for various demographic and geographic domains. Since 1999, it has become a statewide survey that includes 50 States and the District of Columbia. The target population includes civilian, noninstitutionalized persons aged 12 or older. Eight States (California, Florida, Illinois, Michigan, New York, Ohio, Pennsylvania, and Texas), referred to as the "big" States, have a sample designed to yield 3,600 respondents per State, while the remaining 43 "small" States have a sample designed to yield 900 respondents per State. The total sample size is 66,706 persons (corresponding to 51,821 dwelling units [DUs] selected at the second phase out of 169,166 DUs screened at the first phase) with a low of 756 for Nevada to a high of 1,280 for Utah among

"small" States, and a low of 2,669 for New York and a high of 4,681 for California among "big" States.

METHODS: The 1999 NHSDA data for the East South Central Census Division is used to illustrate results obtained by fitting the generalized exponential model (GEM) at both the first phase for nonresponse (nr) and poststratification (ps) of the screener DUs and the second phase for ps for selected persons followed by nr and ps for respondent persons.

RESULTS/CONCLUSIONS: The summary of characteristics of various models fitted is given, as are summary statistics in terms of the unequal weighting effect (UWE), extreme values (ev), and outwinsor proportions, as well as distributional characteristics of the weight distribution. It is seen that % ev is reasonable after ps; therefore, there is no need for an extra ev step. Point estimates and standard errors (SEs) are compared for baseline and final models across a set of drug use variables. For confidentiality reasons, only ratios of point estimates (final over the baseline) are presented; however, individual relative standard errors (RSEs) (SE over the estimate) are shown. Two types of RSE are given; one is unadjusted signifying no adjustment for calibration, and the other is ps-adjusted RSE denoting a sandwich-type formula to adjust for ps, as given in Vaish, Gordek, and Singh (2000) in this volume. It is seen that the two (baseline and final) RSEs are quite comparable. Interestingly, the final RSE can be lower than the baseline RSE although it has more covariates. In cases where the final is higher, it is only marginally so, showing no problems of overfitting.

The generalized exponential model for sampling weight calibration for extreme values, nonresponse, and poststratification

CITATION: Folsom, R. E., & Singh, A. C. (2000). The generalized exponential model for sampling weight calibration for extreme values, nonresponse, and poststratification. In *Proceedings of the 2000 Joint Statistical Meetings, American Statistical Association, Survey Research Methods Section, Indianapolis, IN* (pp. 598-603). Alexandria, VA: American Statistical Association.

PURPOSE/OVERVIEW: There exist methods in the literature that impose bounds on the adjustment factor for poststratification (ps). See, for example, Deville and Särndal (1992), Rao and Singh (1997), and the review by Singh and Mohl (1996). However, they do not directly restrict the adjusted weight from being too extreme. In this paper, the authors consider the problem of developing a unified approach of weight calibration to address the above three concerns such that there are built-in controls on the adjustment factors to prevent the adjusted weight from being too extreme.

METHODS: For this purpose, the logit-type model of Deville and Särndal (1992), denoted by DS in the sequel, is generalized to allow for more general and unit-specific bounds. A review of the DS model is provided in Section 2, and the proposed model is described in Section 3. The asymptotic properties of the proposed calibration estimator are presented in Section 4, and a comparison with alternative methods is given in Section 5. Finally, Section 6 contains numerical results comparing different methods using the 1999 National Household Survey on Drug Abuse (NHSDA) data followed by concluding remarks in Section 7.

RESULTS/CONCLUSIONS: Unlike earlier methods, the generalized exponential model (GEM) provides a unified calibration tool for weight adjustments for extreme values (ev), nonresponse (nr), and ps. Of special interest is its capability to have a built-in control on ev. Under suitable superpopulation modeling and assuming that the bounds on the adjustment factors are prespecified, the resulting calibration estimators were shown to be asymptotically consistent with respect to the distribution for nr and ps, and derivation of the asymptotic Taylor variance estimation formulas analogous to the ones based on residuals for the regression estimator (used for ps) was outlined. In the authors' experience, GEM is a useful practical alternative to the methods of raking-ratio and Deville-Särndal while providing comparable results.

Developing computer-assisted interviewing (CAI) for the National Household Survey on Drug Abuse

CITATION: Lessler, J. T., Caspar, R. A., Penne, M. A., & Barker, P. R. (2000). Developing computer-assisted interviewing (CAI) for the National Household Survey on Drug Abuse. *Journal of Drug Issues, 30*(1), 9-34.

PURPOSE/OVERVIEW: To ensure the privacy of respondents' answers to the National Household Survey on Drug Abuse (NHSDA), the sponsor of the survey in the late 1990s planned to convert it to computer-assisted interviewing (CAI) with an audio computer-assisted self-interviewing (ACASI) section for the most sensitive items. As a result of the survey being self-administered, respondents needed to be able to reconcile inconsistent data during the interview without the assistance of an interviewer.

METHODS: Based on prior findings and research, an ACASI instrument was created that included skip instructions and allowed respondents to reconcile inconsistent answers. A 2x2x2 factorial design was used. The first factor explored whether multigate questions would yield higher prevalence estimates than single-gate questions. The second factor was the ability of respondents to complete data quality consistency checks and its effect on the length of the survey. The third and final factor explored was whether multiple opportunities to report drug use would yield higher estimates. The experiment was embedded in the last quarter of the 1997 NHSDA sample. There were approximately 1,100 respondents in the experiment group and 3,500 in the control group.

RESULTS/CONCLUSIONS: The results showed small but significant time increases for all three factors: multigate questions, inconsistency checks, and multiple opportunities to report drug use. However, the total extra time for these factors was only 1 minute, which was not large enough to rule out elimination on time alone. Contrary to expectations, multigate questions were not found to lead to higher prevalence estimates. Respondents were able to resolve inconsistency checks without unduly increasing the length of the interview. The multiple questions on drug use did not consistently increase or decrease prevalence rates.

Design consistent small area estimates using Gibbs algorithm for logistic models

CITATION: Shah, B. V., Barnwell, B. G., Folsom, R., & Vaish, A. (2000). Design consistent small area estimates using Gibbs algorithm for logistic models. In *Proceedings of the 2000 Joint Statistical Meetings, American Statistical Association, Survey Research Methods Section, Indianapolis, IN* (pp. 105-111). Alexandria, VA: American Statistical Association.

PURPOSE/OVERVIEW: This paper describes the software solutions used to produce age-specific small area estimates and associated pseudo-Bayes posterior intervals for the 50 States and the District of Columbia for the National Household Survey on Drug Abuse (NHSDA).

METHODS: Survey weights were incorporated in conditional posterior distributions to achieve design-consistent asymptotic mean vector and covariance matrices. Proper inverse Wishart priors were used for the covariance matrices. The bias of the estimated fixed effects and covariance components was examined for a three-stage cluster sample design. The software consists of two procedures, PROC GIBBS and PROC GSTAT. PROC GIBBS is used to simulate the posterior distribution for fixed and random effects models using the Gibbs algorithm. PROC GSTAT estimates the prevalence rate of a substance for each block group for each type of individual within the block group for each simulated cycle. Weighted averages of the block group estimates are used to produce county- or State-level estimates.

RESULTS/CONCLUSIONS: The software was able to fit four age-group-specific models simultaneously for over 70,000 observations, 30 to 40 predictors in each model totaling over 100 fixed effects and over 350 random effects. Over 10,000 Gibbs cycles were completed in 10 to 12 hours on a personal computer using the Win95 operating system (300 MHz Pentium II with 256MB RAM). MLwiN and BUGS, two software packages designed to estimate multilevel models, were unable to complete a single cycle. For States with large sample sizes, PROC GIBBS and PROC GSTAT produced estimates that were close to design-based estimates using SUDAAN®.

Bias corrected estimating function approach for variance estimation adjusted for poststratification

CITATION: Singh, A. C., & Folsom, R. E., Jr. (2000). Bias corrected estimating function approach for variance estimation adjusted for poststratification. In *Proceedings of the 2000 Joint Statistical Meetings, American Statistical Association, Survey Research Methods Section, Indianapolis, IN* (pp. 610-615). Alexandria, VA: American Statistical Association.

PURPOSE/OVERVIEW: In this paper, the authors approached the problem of variance estimation adjusted for coverage bias (via poststratification [ps]) by using the conceptual similarity to the variance estimation adjusted for nonresponse (nr) bias, results for which are well known. To this end, it was observed that use of calibration equations for nr and ps adjustments led quite naturally to the nonoptimal estimating function (EF) framework introduced by Binder (1983) for variance estimation for finite population parameters under a quasi-design based framework (i.e., under the joint superpopulation model and design-based distribution), and the

optimal EF framework introduced by Godambe (1960) and Godambe and Thompson (1989) for finite or infinite population parameters. In view of the superpopulation model required for correcting coverage bias or nr or both, the authors therefore proposed a bias-corrected estimating function (BCEF) approach under a quasi-design based framework for estimating variance adjusted for ps or nr or both. The proposed BCEF method was motivated by observing the similarity between ps and nr when one takes the perspective of coverage bias reduction in ps. The BCEF method is based on a simple semiparametric model built on estimating functions that are commonly used for estimating parameters for modeling nr and ps adjustment factor.

METHODS: The authors considered the problem of finding a Taylor linearization variance estimator of the poststratified or the general calibration estimator such that four goals, somewhat analogous to those of Sarndal, Swensson, and Wretman (1989), were met. These goals were (1) the variance estimator should be consistent under a joint design-model distribution (i.e., in a quasi-design based framework), (2) the variance estimator should have a single form with general applicability, (3) the model postulated for the quasi-design based framework should be driven by the real need for unbiased point estimation, and (4) the model should be expected to have sensible conditional properties under a conditional inference outlook whenever it can be suitably defined.

RESULTS/CONCLUSIONS: Using the property inherent to estimating functions, the BCEF method provided a sandwich variance estimate adjusted for ps or nr, or both (this is simply the inverse of the Godambe information matrix), which has a simple yet general form useful for computer automation. Also, using the Godambe finite sample optimality criterion of estimating functions, it was shown that the point and variance estimators (whenever the solution of the estimating function exists) were unique analogous to maximum likelihood estimation for parametric models.

Variance estimation adjusted for weight calibration via the generalized exponential model with application to the National Household Survey on Drug Abuse

CITATION: Vaish, A. K., Gordek, H., & Singh, A. C. (2000). Variance estimation adjusted for weight calibration via the generalized exponential model with application to the National Household Survey on Drug Abuse. In *Proceedings of the 2000 Joint Statistical Meetings, American Statistical Association, Survey Research Methods Section, Indianapolis, IN* (pp. 616-621). Alexandria, VA: American Statistical Association.

PURPOSE/OVERVIEW: In this paper, the authors studied the impact of weight adjustment factors for nonresponse (nr) and poststratification (ps) on the variance of the calibrated sample estimate when adjustment factors were modeled via the generalized exponential model (GEM) of Folsom and Singh (2000) with suitable predictors and bounding restrictions on the adjustment factors.

METHODS: Using the bias-corrected estimating function (BCEF) approach of Singh and Folsom (2000), the estimation problem was cast in the form of estimating equations, which in turn was linearized to obtain a sandwich-type variance estimate of the calibrated estimator. The

method was applied to the 1999 National Household Survey on Drug Abuse (NHSDA), and numerical results comparing variance estimates with and without adjustments were presented.

RESULTS/CONCLUSIONS: The main points are as follows: (1) although the H matrix looks complicated, it followed a nice pattern and could easily be calculated and inverted by using SAS IML; (2) even with a large number of predictor variables used in calibration, the variance estimates adjusted for calibration seemed remarkably stable and, in general, showed gains in efficiency after calibration; and (3) the BCEF methodology was very general and could easily be adapted to other types of calibration techniques. The authors planned to do a validation study by computing resampling variance estimates using Jackknife and comparing the results with the BCEF approach based on the Taylor method. It may be noted that the Jackknife method for obtaining adjusted variance could be quite tedious for large datasets and with somewhat elaborate nr/ps models.

2001

Discussion notes: Session 5

CITATION: Camburn, D., & Hughes, A. (2001). Discussion notes: Session 5. In M. L. Cynamon & R. A. Kulka (Eds.), *Seventh Conference on Health Survey Research Methods* (DHHS Publication No. PHS 01-1013, pp. 251-253). Hyattsville, MD: U.S. Department of Health and Human Services, Public Health Service, Centers for Disease Control and Prevention, National Center for Health Statistics.

PURPOSE/OVERVIEW: The importance of data collection at the State and local level is recognized as a requirement for surveillance and monitoring systems. The presentations in the conference session this paper discusses describe Federal surveillance systems that include as a goal providing State and local area estimates.

METHODS: The session this paper discusses addressed issues affecting the systematic collection of State- and local-level data. The four surveys reported on in the session include the Behavioral Risk Factor Surveillance System (BRFSS), the National Household Survey on Drug Abuse (NHSDA), the National Immunization Survey (NIS), and the State and Local Area Integrated Telephone Survey (SLAITS). The topics discussed include (1) variation in data quality; (2) timeliness of data release; (3) balancing national, State, and local needs; (4) within-State estimates; and (5) analyzing and reporting.

RESULTS/CONCLUSIONS: (1) Concerns about variation in data quality centered on variation in State response rates, cultural differences in the level of respondent cooperation, interviewer effects, and house effects. (2) Survey producers are doing all they can to release the appropriate data in a timely manner, to document limitations, and to minimize microdata disclosure risk. (3) Standardized, well-controlled methodologies may be an advantage in some circumstances, but not others. State-level control over content is important, particularly for within-State analysis, although States are interested in comparability across States. (4) One limitation of current State-level surveillance systems is that resource and time constraints restrict the amount of data that can be collected within individual States for smaller geographic domains or for demographic subgroups. (5) An important issue for surveillance systems collecting data and providing estimates that cover a large number of areas is determining appropriate analysis methods and identifying appropriate methods for reporting the data that include area-specific data quality indicators. For example, in NHSDA, direct estimates are provided for the eight largest States while model-based estimates are calculated for the remaining States. Eventually, NHSDA plans to provide direct estimates for all 50 States.

Variance models applicable to the NHSDA

CITATION: Chromy, J., & Myers, L. (2001). Variance models applicable to the NHSDA. In *Proceedings of the 2001 Joint Statistical Meetings, American Statistical Association, Survey Research Methods Section, Atlanta, GA* [CD-ROM]. Alexandria, VA: American Statistical Association.

PURPOSE/OVERVIEW: When planning or modifying surveys, it is often necessary to project the impact of design changes on the variance of survey estimates. Initial survey planning requires evaluation of the efficiency of alternative sample designs. Efficient designs meet all major requirements at the minimum cost and can only be evaluated in terms of appropriate variance and cost models. This paper focuses on a fairly simple variance model that nevertheless accounts for binomial variation, stratification, intracluster correlation effects, variable cluster sizes, differential sampling of age groups and geographic areas, and residual unequal weighting.

METHODS: The authors examined data from the 1999 National Household Survey on Drug Abuse national sample to develop variance model parameter estimates. They then compared modeled variance estimates with direct estimates of variance based on the application of design-based variance estimation using the SUDAAN® software.

RESULTS/CONCLUSIONS: The modeled relative standard errors provided a realistic approximation to those obtained from design-based estimates. Because they expressed the variance in terms of design parameters, they were useful for evaluating the impact of alternative designs configurations. The simple (unweighted) approach to variance component estimation appeared to provide useful results in spite of ignoring the weights. The impact of unequal weighting was treated as a multiplicative factor. Although the unequal weighting effect can be controlled to some extent by the survey design, the impact of nonresponse and weight adjustment for nonresponse and for calibration against external data can only be controlled in a general way. The unequal weighting effects were not easily subject to any optimization strategy. The authors concluded that the model treatment of variable cluster sizes, particularly for small domains, should be useful in developing variance models for a wide variety of applications.

Estimation for person-pair drug-related characteristics in the presence of pair multiplicities and extreme sampling weights for NHSDA

CITATION: Chromy, J. R., & Singh, A. C. (2001). Estimation for person-pair drug-related characteristics in the presence of pair multiplicities and extreme sampling weights for NHSDA. In *Proceedings of the 2001 Joint Statistical Meetings, American Statistical Association, Survey Research Methods Section, Atlanta, GA* [CD-ROM]. Alexandria, VA: American Statistical Association.

PURPOSE/OVERVIEW: In the National Household Survey on Drug Abuse (NHSDA), Brewer's method is adapted for selecting 0, 1, or 2 persons for the drug survey from the screened dwelling unit (DU) selected at the first phase. Typically, the parameter of interest is at the person level and not the pair level (e.g., at the parent level in the parent-child data). However, pair data are used for estimation because the study variable is measured only through a pair. In estimation, two

major problems arise. One is that of multiplicities because for a given domain, several pairs in a household could be associated with the person, and the multiplicities being domain specific make it difficult to produce a single set of calibrated weights. The other is that of extreme weights due to the possibility of small pair selection probabilities depending on the age groups, which might lead to unstable estimates.

METHODS: For the first problem, the authors propose to do the above calibration simultaneously for controls for key domains. For the second problem, they propose a Hajek-type modification, which entails calibration to controls obtained from larger samples from previous phases. Extreme weights are further reduced by a repeat calibration under bound restrictions while continuing to meet controls.

RESULTS/CONCLUSIONS: It is clear that weights based on pairwise probabilities are required for many drug behavior analyses of the NHSDA data. For this purpose, the analyst needs to make some fundamental decisions about defining population parameters when the person has the same relationship (parent of or child of) to more than one person in the household. For the two problems of multiplicities and extreme weights that might arise in pair data analysis, it was shown how the estimator could be adjusted in the presence of multiplicities and how the weights could be calibrated to alleviate the problem of extreme weights.

Substance use survey data collection methodologies and selected papers [Commentary]

CITATION: Colliver, J., & Hughes, A. (2001). Substance use survey data collection methodologies and selected papers [Commentary]. *Journal of Drug Issues, 31*(3), 717-720.

PURPOSE/OVERVIEW: This paper presents a commentary on the methodological differences between three national surveys that study substance use among adolescents and young adults: the Monitoring the Future (MTF) study, the National Household Survey on Drug Abuse (NHSDA), and the Youth Risk Behavior Survey (YRBS).

METHODS: This paper reviews the current literature discussing the differences in estimates of drug use for the three surveys as a result of differences in documentation, sampling, and survey design.

RESULTS/CONCLUSIONS: The comparative studies sponsored by the Office of Assistant Secretary for Planning and Evaluation provide an excellent resource for understanding the methodological differences in the MTF, the NHSDA, and the YRBS that contribute to the discrepancy in estimates of drug use provided each year.

Coverage, sample design, and weighting in three federal surveys

CITATION: Cowan, C. D. (2001). Coverage, sample design, and weighting in three federal surveys. *Journal of Drug Issues, 31*(3), 599-614.

PURPOSE/OVERVIEW: This paper compares and contrasts coverage and sampling used in three national surveys on drug use among teenagers: the Monitoring the Future (MTF) study, the National Household Survey on Drug Abuse (NHSDA), and the Youth Risk Behavior Survey (YRBS).

METHODS: This review starts by comparing the national teenage population coverage of each of the three surveys to assess the effects of coverage error. Next, the sample design for each study is compared, as well as changes to the methodology made over the years. Finally, the weighting procedures used are analyzed.

RESULTS/CONCLUSIONS: The author concluded that all three studies were well designed, and it was difficult to make recommendations to any of these surveys for improving the assessment of drug abuse among teenagers because it could affect other key variables or subgroups in the individual surveys that are not being compared. However, the author indicated that one recommendation that may not negatively affect the validity of the studies was an in-depth coverage study to assess both frame coverage and nonresponse.

Mode effects in self-reported mental health data

CITATION: Epstein, J. F., Barker, P. R., & Kroutil, L. A. (2001). Mode effects in self-reported mental health data. *Public Opinion Quarterly, 65*(4), 529-549.

PURPOSE/OVERVIEW: This article measures the mode effect differences between audio computer-assisted self-interviewing (ACASI) and an interviewer-administered paper-and-pencil interview (I-PAPI) on respondent reports of mental health issues. Four mental health modules on major depressive episode, generalized anxiety disorder, panic attack, and agoraphobia were taken from the World Health Organization's Composite International Diagnostic Interview (CIDI) short form.

METHODS: The ACASI data were collected in a large-scale field experiment on alternative ACASI versions of the National Household Survey on Drug Abuse (NHSDA). The field experiment was conducted from October through December 1997. The comparison group was comprised of a subsample of the Quarter 4 1997 NHSDA. In the field experiment, mental health questions were administered by ACASI to 865 adults. In the comparison group, mental health questions were administered to a sample of 2,126 adults using I-PAPI. Logistic regression models were used to assess the differences in reporting mental health syndromes by mode of administration. Estimates were made overall and for several demographic variables, such as age, race/ethnicity, gender, education level, geographic region, and population density subgroups while controlling for confounding variables and interactions.

RESULTS/CONCLUSIONS: For most measures, the percentages of people reporting a mental health syndrome were higher for ACASI than I-PAPI. Overall differences were significant only for major depressive episode and generalized anxiety disorder. This study suggests that respondents report higher levels of sensitive behavior with ACASI than with I-PAPI likely due to a perception of greater privacy with ACASI.

Impact of computerized screenings on selection probabilities and response rates in the 1999 NHSDA

CITATION: Eyerman, J., Odom, D., Chromy, J., & Gfroerer, J. (2001). Impact of computerized screenings on selection probabilities and response rates in the 1999 NHSDA. In *Proceedings of the 2001 Joint Statistical Meetings, American Statistical Association, Survey Research Methods Section, Atlanta, GA* [CD-ROM]. Alexandria, VA: American Statistical Association.

PURPOSE/OVERVIEW: The National Household Survey on Drug Abuse (NHSDA) is an ongoing Federal Government survey that tracks substance use in the United States with face-to-face interviews in a national probability sample. In the past, a paper screening instrument was used by field interviewers to identify and select eligible households in the sample frame. The paper screening instrument was replaced with a computerized instrument in 1999. The computerized instrument standardized the screening process and reduced the amount of sampling procedural error in the survey. This paper identifies a type of procedural error that is possible when using paper screening instruments and evaluates its presence in the NHSDA before and after the transition to a computerized screening instrument.

METHODS: The impact of the error is examined against response rates and substance use prevalence estimates for 1999.

RESULTS/CONCLUSIONS: Results suggest that paper screening instruments are vulnerable to sampling procedural error and that the transition to a computerized screening instrument may reduce the amount of the error.

Examining prevalence differences in three national surveys of youth: Impact of consent procedures, mode, and editing rules

CITATION: Fendrich, M., & Johnson, T. P. (2001). Examining prevalence differences in three national surveys of youth: Impact of consent procedures, mode, and editing rules. *Journal of Drug Issues, 31*(3), 615-642.

PURPOSE/OVERVIEW: This review compares contact methods and mode used in three national surveys on drug use conducted in 1997: the Monitoring the Future (MTF) study, the National Household Survey on Drug Abuse (NHSDA), and the Youth Risk Behavior Survey (YRBS).

METHODS: Differences in information presented during the informed consent process are compared for all three studies to evaluate its impact on prevalence estimates of drug abuse. Next, the mode for the three studies was compared focusing on where the study took place (school vs. home environment), when the study took place, and how it was conducted (self-administered vs. interviewer-administered and paper-and-pencil interviewing [PAPI] vs. computer-assisted interviewing [CAI]). Finally, data-editing procedures for inconsistent responses and missing responses were analyzed for each survey.

RESULTS/CONCLUSIONS: Comparisons of these three surveys suggested that the consent process and mode used in the 1997 NHSDA may have contributed to the lower prevalence estimates compared with the other 1997 studies. The NHSDA consent process required more parental involvement and presented more consent information than did the process used by the other two studies, which may have inhibited respondent reporting. Differences in editing procedures did not appear to account for any differences in prevalence estimates for the three surveys.

Learning from experience: Estimating teen use of alcohol, cigarettes, and marijuana from three survey protocols

CITATION: Fowler, F. J., Jr., & Stringfellow, V. L. (2001). Learning from experience: Estimating teen use of alcohol, cigarettes, and marijuana from three survey protocols. *Journal of Drug Issues, 31*(3), 643-664.

PURPOSE/OVERVIEW: This paper compares prevalence estimates for drug use among teenagers in the three national surveys funded by the Federal Government: the Monitoring the Future (MTF) study, the National Household Survey on Drug Abuse (NHSDA), and the Youth Risk Behavior Survey (YRBS).

METHODS: Because the three surveys rely on different modes for data collection, comparisons had to be made across similar groups. For that reason, rates at which adolescents in grades 10 and 12 reported drug use were compared. In addition, comparisons were made between males and females and across racial/ethnic groups, such as whites, Hispanics, and blacks. Finally, trends in drug use from 1993 to 1997 were compared across surveys.

RESULTS/CONCLUSIONS: Comparisons across these surveys were hard to draw because any differences in prevalence estimates were due to numerous confounding factors associated with coverage, sampling, mode, questionnaires, and data collection policies. A solution suggested is for each of these studies to set aside a small amount of money that can be used for collaborative studies of comparisons of the methods used.

Substance use survey data collection methodologies and selected papers [Commentary]

CITATION: Gfroerer, J. (2001). Substance use survey data collection methodologies and selected papers [Commentary]. *Journal of Drug Issues, 31*(3), 721-724.

PURPOSE/OVERVIEW: The Office of the Assistant Secretary for Planning and Evaluation sponsored five comparative papers to assess differences in methodologies used by three national surveys on drug use among adolescents.

METHODS: The papers compared the National Household Survey on Drug Abuse (NHSDA), the Monitoring the Future (MTF) study, and the Youth Risk Behavior Survey (YRBS).

Differences in sampling, mode, and other data collection properties were assessed to determine their impact on the differences in prevalence estimates from the three surveys.

RESULTS/CONCLUSIONS: Finding differences in estimates can provide information to inform an understanding of the current estimates. The findings of these papers also will help target specific areas for more methodological research. Although these studies were sponsored to explain differences in estimates between the surveys, it also is important to note several of the consistencies in these surveys, such as in demographic differences and trends over time.

State estimates of substance abuse prevalence: Redesign of the National Household Survey on Drug Abuse (NHSDA)

CITATION: Gfroerer, J., Wright, D., & Barker, P. (2001). State estimates of substance abuse prevalence: Redesign of the National Household Survey on Drug Abuse (NHSDA). In M. L. Cynamon & R. A. Kulka (Eds.), *Seventh Conference on Health Survey Research Methods* (DHHS Publication No. PHS 01-1013, pp. 227-232). Hyattsville, MD: U.S. Department of Health and Human Services, Public Health Service, Centers for Disease Control and Prevention, National Center for Health Statistics.

PURPOSE/OVERVIEW: Starting with the 1999 survey, the National Household Survey on Drug Abuse (NHSDA) sample was expanded and redesigned to improve its capability to estimate substance use prevalence in all 50 States. This paper discusses the new NHSDA design, data generated from the new design, and issues related to its implementation. Also provided are a summary of the old survey design and a discussion of other major changes implemented in 1999, such as the conversion of the survey from paper-and-pencil interviewing (PAPI) to computer-assisted interviewing (CAI).

METHODS: The pre-1999 survey design and the new survey design are compared and discussed in terms of various methodological implications: (1) data collection methodology, (2) questionnaire content, (3) data limitations, (4) sample design, (5) issues in implementing the expanded sample, (6) estimates that the new design will provide, and (7) estimation and analysis issues.

RESULTS/CONCLUSIONS: The findings include the following: (1) Although the basic research methodology has remained unchanged, the conversion from PAPI to CAI incorporated features designed to decrease burden and increase data quality, including electronic screening, range edits, and consistency checks throughout the questionnaire, and inconsistency resolution in audio computer-assisted self-interviewing (ACASI). (2) Although the content of the 1999 CAI questionnaire is similar to the 1998 PAPI questionnaire, the CAI interview length is considerably shorter than the PAPI interview length. (3) Although the methods used in the survey have been shown effective in reducing reporting bias, there is still probably some unknown level of underreporting that occurs. (4) In response to a need for comparable State-level estimates of substance abuse prevalence, in 1999 SAMHSA expanded the survey sample. This was determined to be feasible based on prior experiences with modeling for selected States, as well as a sampling plan for 1999 that would facilitate State-level estimation. (5) The size and distribution of the 1999 sample units across the 50 States posed a challenge for data collection

operations. In spite of extensive training, the inexperience of new interviewers hired for the 1999 survey expansion led to a decline in response rates relative to prior NHSDAs. (6) The sample was designed to produce both model-based and sample-based State-level estimates of a variety of substance use measures. (7) Estimation and analysis issues include comparability of NHSDA State estimates, comparisons of NHSDA State-level estimates to other surveys in States, assessing trends within States, and data release and disclosure limitation.

Understanding the differences in youth drug prevalence rates produced by the MTF, NHSDA, and YRBS studies

CITATION: Harrison, L. D. (2001). Understanding the differences in youth drug prevalence rates produced by the MTF, NHSDA, and YRBS studies. *Journal of Drug Issues, 31*(3), 665-694.

PURPOSE/OVERVIEW: This paper explores methodological differences in three national surveys of drug use in order to explain the discrepancy in prevalence estimates between the three sources.

METHODS: The three surveys examined in this paper are the National Household Survey on Drug Abuse (NHSDA), the Monitoring the Future (MTF) study, and the Youth Risk Behavior Survey (YRBS). This paper explores the validity of the survey estimates provided by these studies by comparing their differences in methodology, such as survey design, anonymity, confidentiality, and question context.

RESULTS/CONCLUSIONS: Despite the purpose of this paper being to explore the differences in estimates in these three surveys, analyses revealed that the estimates might be more similar than originally thought. The confidence intervals for several variables in the surveys overlapped, and trend analyses in drug use for the surveys followed the same pattern. Any differences found between the survey estimates were likely a result of many minor methodological differences between the surveys. No one survey can be shown to be more accurate than another, and in fact using several different sources provides a more informed overall picture of drug use among youths. The author concluded that each survey should continue using its current methodology because although different, each fulfills a particular need. However, more studies could be done to continue learning about the impact of methodological differences in each survey.

Substance use survey data collection methodologies: Introduction

CITATION: Hennessy, K. H., & Ginsberg, C. (Eds.). (2001). Substance use survey data collection methodologies: Introduction. *Journal of Drug Issues, 31*(3), 595-598.

PURPOSE/OVERVIEW: There are three annual surveys funded by the U.S. Department of Health and Human Services that collect data on the prevalence of substance use and abuse among youths. The results of these studies are used to form programs and influence policies to address the substance use problem among youths.

METHODS: The Office of the Assistant Secretary for Planning and Evaluation was tasked with comparing and contrasting the survey design, sampling, and statistical assessment used in the three surveys: the National Household Survey on Drug Abuse (NHSDA), the Monitoring the Future (MTF) study, and the Youth Risk Behavior Survey (YRBS).

RESULTS/CONCLUSIONS: Although the three surveys produce different prevalence estimates for drug use and abuse among youths, all three have strong designs and have shown similar trends over time. The papers commissioned to compare the methodologies of these three studies should further aid in understanding the differences between these surveys and assist policymakers in tracking the prevalence and trends of drug use.

Substance use survey data collection methodologies [Commentary]

CITATION: Kann, L. (2001). Substance use survey data collection methodologies [Commentary]. *Journal of Drug Issues, 31*(3), 725-728.

PURPOSE/OVERVIEW: Alcohol and drug use is a leading social and health concern in the United States. It is associated with and contributes to higher mortality rates and crime. Therefore, a national surveillance of alcohol and drug use is an integral public health issue.

METHODS: This paper is a commentary on the five papers sponsored by the Office of the Assistant Secretary for Planning and Evaluation to evaluate the differences in methodology used in the leading surveys on drug use in the United States and to assess the impact on prevalence estimates for drug use.

RESULTS/CONCLUSIONS: The author concluded that the five papers were a solid resource in understanding the different methodologies used in collecting information about drug use. The papers also highlighted further methodological work that could be done to clarify more fully the effects of methodological differences on the prevalence estimates for the three surveys.

Needs for state and local data of national relevance

CITATION: Lepkowski, J. M. (2001). Needs for state and local data of national relevance. In M. L. Cynamon & R. A Kulka (Eds.), *Seventh Conference on Health Survey Research Methods* (DHHS Publication No. PHS 01-1013, pp. 247-250). Hyattsville, MD: U.S. Department of Health and Human Services, Public Health Service, Centers for Disease Control and Prevention, National Center for Health Statistics.

PURPOSE/OVERVIEW: Data from health surveys provide information critical to State and local governments for health policy and resource allocation decisions. The four surveys described in the session that this discussion paper addresses employ various methodological approaches for collecting health data for State and local areas.

METHODS: The four surveys addressed in this discussion paper include: the Behavioral Risk Factor Surveillance System (BRFSS), the National Household Survey on Drug Abuse (NHSDA),

the National Immunization Survey (NIS), and the State and Local Area Integrated Telephone Survey (SLAITS).

RESULTS/CONCLUSIONS: The four surveys share important design goals, such as producing valid estimates for the entire country and for States. However, the BRFSS, NIS, and SLAITS all have noncoverage concerns because of the use of telephone sampling methods for the household portion of the population. Also, although nonresponse is present in all four surveys, the potential nonresponse bias is greatest for the three surveys employing telephone sampling methods. The combination of noncoverage error and nonresponse error can be problematic, especially at the State level. The author suggests several ways to address this concern, including compensatory weights for noncoverage and nonresponse, the use of supplemental frames, and researching models to fully address factors that affect nonresponse. The author also notes that improved survey content and the coordination of State- and nationally administered surveys are of great importance to increasing the quality of State and local health surveys.

Development of computer-assisted interviewing procedures for the National Household Survey on Drug Abuse

CITATION: Office of Applied Studies. (2001, March). *Development of computer-assisted interviewing procedures for the National Household Survey on Drug Abuse* (DHHS Publication No. SMA 01-3514, Methodology Series M-3). Rockville, MD: Substance Abuse and Mental Health Services Administration.

PURPOSE/OVERVIEW: In 1999, the National Household Survey on Drug Abuse (NHSDA) sample was expanded and redesigned to permit using a combination of direct and model-based small area estimation (SAE) procedures that allow the Substance Abuse and Mental Health Services Administration (SAMHSA) to produce estimates for all 50 States and the District of Columbia. In addition, computer-assisted data collection procedures were adopted for both screening and interviewing. This report summarizes the research to develop these computer-assisted screening and interviewing procedures.

METHODS: This report covers a variety of NHSDA field experiment topics. To start, Chapter 2 gives a brief history of research on the NHSDA, and Chapter 3 offers further background information, including a literature review and an overview of critical design and operational issues. Chapter 4 focuses on the 1996 feasibility experiment and cognitive laboratory research, while Chapters 5 through 9 delve into the 1997 field experiment. Specifically, Chapter 5 summarizes the design and conduct of the 1997 effort, Chapter 6 compares computer-assisted personal interviewing (CAPI) and audio computer-assisted self-interviewing (ACASI) with paper-and-pencil interviewing (PAPI) for selected outcomes, Chapter 7 describes the effect of ACASI experimental factors on prevalence and data quality, Chapter 8 details the development and testing of an electronic screener, and Chapter 9 describes the operation of the 1997 field experiment. The next two chapters offer insights into the willingness of NHSDA respondents to be interviewed (Chapter 10) and the effect of NHSDA interviewers on data quality (Chapter 11). Chapter 12 is devoted to further refinement of the computer-assisted interviewing (CAI) procedures during the 1998 laboratory and field testing of a tobacco module.

RESULTS/CONCLUSIONS: The main results from the 1997 field experiment were the following: (1) Use of a single gate question to ask about substance use rather than multiple gate questions yielded higher reporting of substance use. (2) The use of consistency checks within the survey instrument yielded somewhat higher reporting of drug use than when such checks were not used. (3) The use of ACASI yielded higher reports of drug use than PAPI. (4) Respondents were less likely to request help in completing the survey in ACASI than in PAPI, particularly among youths (aged 12–17) and adults with less than a high school education. (5) Respondents using ACASI reported higher comfort levels with the interview than those using PAPI. (6) Respondents with fair or poor reading ability found the recorded voice in ACASI more beneficial than those with excellent reading ability. (7) ACASI respondents reported higher levels of privacy than PAPI respondents.

National Household Survey on Drug Abuse: 1999 nonresponse analysis report

CITATION: Office of Applied Studies. (2001). *National Household Survey on Drug Abuse: 1999 nonresponse analysis report*. Rockville, MD: Substance Abuse and Mental Health Services Administration.

PURPOSE/OVERVIEW: This report addresses the nonresponse patterns obtained in the 1999 National Household Survey on Drug Abuse (NHSDA). This report was motivated by the relatively low response rates in the 1999 NHSDA and by the apparent general trend of declining response rates in field studies. The analyses presented in this report were produced to help provide an explanation for the rates in the 1999 NHSDA and guidance for the management of future projects.

METHODS: The six chapters of this report provide a background for the issues surrounding nonresponse and an analysis of the 1999 NHSDA nonresponse. The first three chapters provide context with reviews of the nonresponse trends in U.S. field studies, the current academic literature, and the NHSDA data collection patterns from 1994 through 1998. Chapter 4 describes the data collection process in 1999 with a detailed discussion of design changes, summary figures and statistics, and a series of logistic regressions. Chapter 5 compares 1998 with 1999 nonresponse patterns. Chapter 6 applies the analysis in the previous chapters to a discussion of a respondent incentive program for future NHSDA work.

RESULTS/CONCLUSIONS: The results of this study are consistent with the conventional wisdom of the professional survey research field and the findings in survey research literature. The nonresponse can be attributed to a set of interviewer influences, respondent influences, design features, and environmental characteristics. The nonresponse followed the demographic patterns observed in other studies, with urban and high crime areas having the worst rates. Finally, efforts taken to improve the response rates were effective. Unfortunately, the tight labor market combined with the large increase in sample size caused these efforts to lag behind the data collection calendar. The authors used the results to generate several suggestions for the management of future projects. First, efforts should be taken to convert reluctant sample elements to completions. This report contains an outline for an incentive program that addresses this issue. Second, because the characteristics of field staff are among the most important correlates of nonresponse, a detailed analysis should be conducted to evaluate the most effective

designs for staffing and retention. Finally, actions should be taken to tailor the survey to regional characteristics, such as environmental and respondent characteristics, which are important predictors of response patterns.

Culture and item nonresponse in health surveys

CITATION: Owens, L., Johnson, T. P., & O'Rourke, D. (2001). Culture and item nonresponse in health surveys. In M. L. Cynamon & R. A Kulka (Eds.), *Seventh Conference on Health Survey Research Methods* (DHHS Publication No. PHS 01-1013, pp. 69-74). Hyattsville, MD: U.S. Department of Health and Human Services, Public Health Service, Centers for Disease Control and Prevention, National Center for Health Statistics.

PURPOSE/OVERVIEW: Item nonresponse is one of the available indicators that can be used to identify cultural variations in survey question interpretation and response. This paper analyzes the patterns of item nonresponse across cultural groups in the United States using data from four national health surveys. The authors hypothesized that respondents from minority cultural groups would exhibit higher nonresponse to survey questions than non-Hispanic white respondents.

METHODS: For the analysis, the authors used four national health-related surveys: the 1992 Behavioral Risk Factor Surveillance System (BRFSS), the 1992 National Household Survey on Drug Abuse (NHSDA), the 1991 National Health Interview Drug Use Supplement (NHIS), and the 1990–1991 National Comorbidity Survey (NCS). In each dataset, the authors selected several items that reflected different health domains. From these items, the authors created 10 dichotomous variables that measured whether the source items contained missing data. Each measure was examined using simple cross-tabulations and logistic regression models in which the authors controlled for sociodemographic variables associated with item nonresponse (e.g., age, gender, education, and marital status).

RESULTS/CONCLUSIONS: Although item nonresponse rates of health survey questions appeared to be low, the authors found that item nonresponse may vary systematically across cultural groups. The authors also found higher item nonresponse rates among each of the minority racial and ethnic groups as compared with non-Hispanic white respondents. In addition to these general findings, the authors noted group-specific cultural differences that warrant special attention. First, the largest odds ratio associated with respondent culture, which reflected Hispanic refusals to answer questions related to their social relationships, may be a consequence of a cultural difference: showing unwillingness to report anything other than positive relations with their family and friends. Second, greater reluctance to report substance use by African-American respondents may be explained by this group's beliefs about selective enforcement of drugs laws against minority groups in the United States. Finally, although the generally low prevalence of item nonresponse in the data analyzed may indicate that differential rates of nonresponse are of insignificant magnitude to seriously bias survey findings, the data do suggest that cultural differences in item nonresponse may be more problematic under certain conditions.

Person-pair sampling weight calibration using the generalized exponential model for the National Household Survey on Drug Abuse

CITATION: Penne, M. A., Chen, P., & Singh, A. C. (2001). Person-pair sampling weight calibration using the generalized exponential model for the National Household Survey on Drug Abuse. In *Proceedings of the 2001 Joint Statistical Meetings, American Statistical Association, Social Statistics Section, Atlanta, GA* [CD-ROM]. Alexandria, VA: American Statistical Association.

PURPOSE/OVERVIEW: For pair data analysis, sampling weights need to be calibrated such that the problems of extreme weights are addressed. Extreme weights, which give rise to high unequal weighting effect (UWE) and hence instability in estimation, are derived from possible small pairwise selection probabilities and domain-specific multiplicity factors.

METHODS: The recently developed generalized exponential model (GEM) for weight calibration at RTI by Folsom and Singh (2000) allows for multiple calibration controls, as well as separate bounds on the weight adjustment factors for extreme weights identified before calibration. Thus, controls corresponding to the number of pairs in various demographic groups from the first phase of screened dwelling units, and controls for the number of individuals in the domains of interest from the second phase of surveyed pairs and single persons, for a key set of domains, can all be incorporated simultaneously in calibration, giving rise to a final single set of calibration weights. Numerical examples of GEM calibration of pair data and the resulting estimates for the 1999 National Household Survey on Drug Abuse (NHSDA) are presented.

RESULTS/CONCLUSIONS: Both model groups of region pair samples of the 1999 NHSDA are used to illustrate results obtained by utilizing the GEM methodology to calibrate a final analytic weight. The authors present summary results of before and after each pair adjustment step in the calibration process for the Northeast and South regions and the Midwest and West regions, respectively. Note that "after" results of each preceding adjustment step is synonymous with "before" results of every subsequent step. Sample sizes; the UWE; unweighted, weighted, and outwinsor percentages of extreme values; and the quartile distribution of both the weight component itself and the weight product up through that step also are presented.

How do response problems affect survey measurement of trends in drug use?

CITATION: Pepper, J. V. (2001). How do response problems affect survey measurement of trends in drug use? In S. K. Goldsmith, T. C. Pellmar, A. M. Kleinman, W. E. Bunney, & Commission on Behavioral and Social Sciences and Education (Eds.), *Informing America's policy on illegal drugs: What we don't know keeps hurting us* (pp. 321-348). Washington, DC: National Academy Press.

PURPOSE/OVERVIEW: Two databases are widely used to monitor the prevalence of drug use in the United States. The Monitoring the Future (MTF) study surveys high school students, and the National Household Survey on Drug Abuse (NHSDA) surveys the noninstitutionalized residential population aged 12 or older. Each year, respondents from these surveys are drawn from known populations—students and noninstitutionalized people—according to well-specified

probabilistic sampling schemes. Hence, in principle, these data can be used to draw statistical inferences on the fractions of the surveyed populations who use drugs. It is inevitable, however, for questions to be raised about the quality of self-reports of drug use. Two well-known response problems hinder one's ability to monitor levels and trends: nonresponse, which occurs when some members of the surveyed population do not respond, and inaccurate response, which occurs when some surveyed persons give incorrect responses to the questions posed. These response problems occur to some degree in almost all surveys. In surveys of illicit activity, however, there is more reason to be concerned that decisions to respond truthfully, if at all, are motivated by respondents' reluctance to report that they engage in illegal and socially unacceptable behavior. To the extent that nonresponse and inaccurate response are systematic, surveys may yield invalid inferences about illicit drug use in the United States.

METHODS: To illustrate the inferential problems that arise from nonresponse and inaccurate response, the author suggested using the MTF and the NHSDA to draw inferences on the annual prevalence of use for adolescents. Annual prevalence measures indicate use of marijuana, cocaine, inhalants, hallucinogens, heroin, or nonmedical use of psychotherapeutics at least once during the year. Different conclusions about levels and trends might be drawn for other outcome indicators and for other subpopulations.

RESULTS/CONCLUSIONS: The author concluded that the MTF and the NHSDA provide important data for tracking the numbers and characteristics of illegal drug users in the United States. Response problems, however, continued to hinder credible inference. Although nonresponse may have been problematic, the lack of detailed information on the accuracy of response in the two national drug use surveys was especially troubling. Data were not available on the extent of inaccurate reporting or on how inaccurate response changes over time. In the absence of good information on inaccurate reporting over time, inferences on the levels and trends in the fraction of users over time were largely speculative. It might be, as many had suggested, that misreporting rates were stable over time. It also might have been that these rates varied widely from one period to the next. The author indicated that these problems, however, did not imply that the data were uninformative or that the surveys should be discontinued. Rather, researchers using these data must either tolerate a certain degree of ambiguity or must be willing to impose strong assumptions. The author suggested practical solutions to this quandary: If stronger assumptions are not imposed, the way to resolve an indeterminate finding is to collect richer data. Data on the nature of the nonresponse problem (e.g., the prevalence estimate of nonrespondents) and on the nature and extent of inaccurate response in the national surveys might be used to both supplement the existing data and to impose more credible assumptions. Efforts to increase the valid response rate may reduce the potential effects of these problems.

Predictive mean neighborhood imputation with application to the person-pair data of the National Household Survey on Drug Abuse

CITATION: Singh, A., Grau, E., & Folsom, R., Jr. (2001). Predictive mean neighborhood imputation with application to the person-pair data of the National Household Survey on Drug Abuse. In *Proceedings of the 2001 Joint Statistical Meetings, American Statistical Association, Survey Research Methods Section, Atlanta, GA* [CD-ROM]. Alexandria, VA: American Statistical Association.

PURPOSE/OVERVIEW: The authors present a simple method of imputation termed "predictive mean neighborhood," or PMN. Features of this method include the following: (1) it allows for several covariates; (2) the relative importance of covariates is determined objectively based on their relationship to the response variable; (3) it incorporates design weights; (4) it can be multivariate in that correlations between several imputed variables are preserved, as well as correlations across imputed and observed variables; (5) it accommodates both discrete and continuous variables for multivariate imputation; and finally (6) it should lend itself to a simple variance estimation adjusted for imputation.

METHODS: The PMN method is a combination of prediction modeling with a random nearest neighbor hot deck. It uses a model-based predictive mean to find a small neighborhood of donors from which a single donor is selected at random. Thus, the residual distribution is estimated nonparametrically from the donors, where the imputed value is (approximately) the predictive mean plus a random residual. Applications of PMN for imputing two types of multiplicity factors required for pair data analysis from the 1999 NHSDA are discussed.

RESULTS/CONCLUSIONS: The PMN methodology has been widely used for the imputation of a variety of variables in the NHSDA, including both continuous and categorical variables with one or more levels. The models were fit using standard modeling procedures in SAS® and SUDAAN®, while SAS macros were used to implement the hot-deck step, including the restrictions on the neighborhoods. Although creating a different neighborhood for each item nonrespondent was computationally intensive, the method was implemented successfully. At the time this paper was presented, the imputations team at RTI was implementing a series of simulations to evaluate the new method, comparing it against the unweighted sequential hot deck used earlier and a simpler model-based method.

Examining substance abuse data collection methodologies

CITATION: Sudman, S. (2001). Examining substance abuse data collection methodologies. *Journal of Drug Issues, 31*(3), 695-716.

PURPOSE/OVERVIEW: The U.S. Department of Health and Human Services (DHHS) sponsors three national surveys on drug use in adolescents. These three studies are the National Household Survey on Drug Abuse (NHSDA), the Youth Risk Behavior Surveillance System (YRBSS), and the Monitoring the Future (MTF) study. The estimates for drug use reported for three surveys differ considerably. This paper examines methodological reasons for the differences reported in these studies.

METHODS: The author does not conduct any new experiments or analyses on the survey data, but examines previous validation and methodological research. The major differences in methodology identified, which might contribute to differences in estimates, are mode of administration (home vs. school setting), questionnaire context and wording, sample design, and weighting.

RESULTS/CONCLUSIONS: Of all the methodological differences, it appears that the context and introduction to the questionnaires in addition to the mode of administration have the largest

impact because they affect the respondents' perceived anonymity. It is likely that NHSDA respondents do not perceive the study to be as anonymous as respondents in both the MTF and YRBSS. More research should be conducted to ascertain how anonymous respondents felt the study was.

2002

Assessment of the computer-assisted instrument

CITATION: Caspar, R., & Penne, M. (2002). Assessment of the computer-assisted instrument. In J. Gfroerer, J. Eyerman, & J. Chromy (Eds.), *Redesigning an ongoing national household survey: Methodological issues* (DHHS Publication No. SMA 03-3768, pp. 53-84). Rockville, MD: Substance Abuse and Mental Health Services Administration, Office of Applied Studies.

PURPOSE/OVERVIEW: The conversion of the National Household Survey on Drug Abuse (NHSDA) to computer-assisted interviewing (CAI) offered an opportunity to improve the quality of the data collected in the NHSDA in a number of ways. Some of these improvements were implemented easily and manifested themselves in more complete data (e.g., the ability to eliminate situations where questions were inadvertently left blank by the respondent). However, other improvements could only be realized through careful development and implementation of new procedures. Thorough testing was needed to determine whether these new procedures did, in fact, result in higher quality data.

METHODS: This chapter describes two significant revisions to how key NHSDA data items are collected and the effect of these revisions on the quality of the data obtained in the 1999 NHSDA. The first of these revisions was the addition of a methodology for resolving inconsistent or unusual answers provided by the respondent. This methodology was incorporated into the collection of a large number of the data items that are considered critical to the reporting needs of the NHSDA. The second revision dealt specifically with the way data on frequency of substance use over the past 12-month period was reported. This chapter also provides a review of several basic measures of data quality, including rates of "Don't Know" and "Refused" responses, breakoff interviews, and the observational data provided by the interviewers at the conclusion of each interview. Where possible, these measures are compared between the CAI and paper-and-pencil interviewing (PAPI) NHSDA instruments as a means of assessing the effect of the move to CAI on data quality.

RESULTS/CONCLUSIONS: The results suggested that the move to CAI data collection has improved data quality, although in some cases the increase was fairly small because data quality for the PAPI NHSDA data was already quite high. Perhaps the most significant improvement to data quality came as a result of the inclusion of the inconsistent and unusual data checks (described in Section 4.1). This enhancement was a radical departure from the PAPI NHSDA and one that was possible only under the CAI mode of interview. The results presented here show that, although the CAI data did not suffer from a large amount of inconsistent or unusual data, the methodology was able to resolve a large number of these cases in a way that was both cost-effective and likely to enhance overall data quality for the items involved. Results of the change in the 12-month frequency of use item were somewhat difficult to interpret. The revised method for collecting these data resulted in higher reported frequencies, but whether this was due to the revision or simply the move from PAPI to CAI was impossible to determine. The distribution of responses and the fact that the mean frequency was higher under CAI than PAPI provided anecdotal support for the revised method of collecting these data, however. Finally,

results from basic measures of data quality and interviewer debriefing items suggested that the CAI methodology reduced interview difficulties among respondents, helped to further enhance the degree of privacy, and appeared to contribute positively to item-level response rates.

Mode effects on substance use measures: Comparison of 1999 CAI and PAPI data

CITATION: Chromy, J., Davis, T., Packer, L., & Gfroerer, J. (2002). Mode effects on substance use measures: Comparison of 1999 CAI and PAPI data. In J. Gfroerer, J. Eyerman, & J. Chromy (Eds.), *Redesigning an ongoing national household survey: Methodological issues* (DHHS Publication No. SMA 03-3768, pp. 135-159). Rockville, MD: Substance Abuse and Mental Health Services Administration, Office of Applied Studies.

PURPOSE/OVERVIEW: The shift from paper-and-pencil interviewing (PAPI) to computer-assisted interviewing (CAI) in the 1999 National Household Survey on Drug Abuse (NHSDA) was anticipated to have significant effects on the reporting of substance use by survey respondents. This expectation was based on several studies in the literature, as well as field testing done within the NHSDA project that showed respondents were more willing to report sensitive behaviors using audio computer-assisted interviewing than with self-administered paper questionnaires. Because of the great interest in analyzing trends in substance use prevalence, a critical component of the 1999 NHSDA redesign was the supplemental sample that employed the "old" PAPI NHSDA data collection methodology. The intent of this dual-sample design in 1999 was primarily to make it possible to continue to measure trends, using estimates from both before and after the redesign. However, the large dual-sample design of the 1999 NHSDA also can be viewed as an important survey research experiment assessing mode effects on the reporting of sensitive behaviors. With its dual-sample design, the 1999 NHSDA provided a large sample for assessing the impact of mode of interview. However, the intention of the 1999 NHSDA design was not to evaluate the mode effect, but to determine the impact of the overall change in method due to the redesign of the survey in 1999 on the time series of substance use statistics. The overall change involved many aspects of the survey design and estimation procedures in addition to mode, such as the sampling plan, the questionnaire, data editing, and imputation. Isolating the "pure" mode effect with these data was difficult and was complicated further by the unexpected impact of interviewer experience on substance use prevalence estimates (see Chapter 8 of the 2002 report). Nevertheless, the analyses presented in this chapter provide some important findings concerning mode effects, as well as on the comparability of pre-1999 NHSDA published estimates with estimates from the redesigned NHSDA. When studying the reporting of sensitive or illegal behaviors, conventional thinking has been that higher reporting is closer to the truth. This evaluation continued that approach, recognizing that this may not be true in every case.

METHODS: The authors compare the substance use prevalence estimates derived from the 1999 PAPI and CAI samples. In addition to providing NHSDA data users with information that will help them interpret NHSDA trends in substance use prevalence, the analysis also is of interest to survey researchers concerned with mode effects.

RESULTS/CONCLUSIONS: The results support previous research that shows higher reporting of sensitive behaviors with audio computer-assisted self-interviewing (ACASI) than with self-administered paper answer sheets. A total of 336 comparisons of unedited estimates from PAPI and CAI were made. Of these, 112 indicated significantly higher CAI estimates, while only 5 indicated significantly higher PAPI estimates. Higher CAI estimates were particularly evident for lifetime prevalence, the measures impacted least by questionnaire structure differences between CAI and PAPI. The analyses of edited and imputed NHSDA estimates showed mixed results for past year and past month use. Lifetime prevalence rates, which were minimally affected by editing and imputation, were generally higher with CAI than with PAPI. A total of 448 comparisons of edited and imputed estimates from PAPI and CAI were made. Of a total of 62 statistically significant differences in lifetime prevalence, 56 indicated a higher CAI estimate and 6 indicated a higher PAPI estimate. Results for past year and past month measures showed variation across substances and age groups and also were different in the full sample analysis than in the matched sample analysis. Out of 87 significant differences for past month or past year use, 41 indicated a higher CAI estimate and 46 indicated a higher PAPI estimate. For past month use of alcohol and cigarettes, the substances with the highest prevalence of use, the observed PAPI estimate was higher than the CAI estimate for every age group and for both the full and matched sample analyses, and 11 of the 16 comparisons were statistically significant. Marijuana use estimates also tended to be higher with PAPI, although most of these differences were not statistically significant. One clear conclusion is that the CAI mode of interviewing led to more internally consistent and complete data. Under PAPI, the need for editing and imputation to clarify recency of use was larger for most substances and provided a greater opportunity to influence estimates through the editing and imputation process. In summary, the CAI methodology produced more complete data with a lower requirement for editing and higher prevalence estimates when treating only unambiguous reports as positive indications of substance use.

Nonresponse in the 1999 NHSDA

CITATION: Eyerman, J., Odom, D., Wu, S., & Butler, D. (2002). Nonresponse in the 1999 NHSDA. In J. Gfroerer, J. Eyerman, & J. Chromy (Eds.), *Redesigning an ongoing national household survey: Methodological issues* (DHHS Publication No. SMA 03-3768, pp. 23-51). Rockville, MD: Substance Abuse and Mental Health Services Administration, Office of Applied Studies.

PURPOSE/OVERVIEW: The redesign of the National Household Survey on Drug Abuse (NHSDA) in 1999 resulted in major changes in many aspects of the data collection procedures. This raised concerns that the response rates could be affected. In particular, the increased sample size, reduced clustering of sample segments, transition from a paper to a computerized screening instrument, and the transition from paper-and-pencil interviewing (PAPI) to computer-assisted interviewing (CAI) all had the potential to change the response rates. During the first quarter of 1999, an assessment of the progress in completing the fieldwork indicated a reduction in the response rates, relative to response rates achieved historically in the NHSDA. To address this problem, several management actions were implemented immediately. Although the response rates improved steadily throughout the remainder of the year, the result was significantly lower response rates for the 1999 NHSDA than for prior NHSDAs.

METHODS: Extensive analysis was undertaken in an attempt to understand the reasons for the drop in response rates and how it was related to each of the design changes. This chapter summarizes this analysis. It also discusses the management actions implemented during 1999 to improve the response rates and assesses the effectiveness of these actions.

RESULTS/CONCLUSIONS: The design changes between 1998 and 1999 corresponded with a large decrease in the response rates. A series of management efforts was taken to address the decrease, both in anticipation of the design changes and in reaction to unexpected results of the changes. In general, the efforts were successful. The extended analysis presented in this chapter summarizes the current understanding of the decline in the response rates in 1999. First, it appears that the previous understanding of the correlates of nonresponse was correct, but it does not completely explain the difference between 1998 and 1999. Second, management efforts taken during 1999 appear to have been successful in reducing the decline in the response rates, and this success carried over to 2000. Third, the computerized screening instrument reduced sampling bias by removing interviewer effects from the screening routine, and this had a small and negative impact on the response rates. Fourth, the transition from PAPI to CAI increased the response rates. Finally, much of the decline in 1999 can be attributed to changes in the composition of the field staff resulting from the large increase in sample size, most notably the reduced number of experienced field interviewers working on the project. However, this does not fully explain the decline in 1999.

Introduction. In J. Gfroerer, J. Eyerman, & J. Chromy (Eds.), Redesigning an ongoing national household survey: Methodological issues

CITATION: Gfroerer, J. (2002). Introduction. In J. Gfroerer, J. Eyerman, & J. Chromy (Eds.), *Redesigning an ongoing national household survey: Methodological issues* (DHHS Publication No. SMA 03-3768, pp. 1-8). Rockville, MD: Substance Abuse and Mental Health Services Administration, Office of Applied Studies.

PURPOSE/OVERVIEW: In 1999, a major redesign of the National Household Survey on Drug Abuse (NHSDA) was implemented involving both the sample design and the data collection method of the survey. The strictly national design was changed to a much larger, State-based design to meet the needs of policymakers for estimates of substance use prevalence for each State. The data collection method was changed from a paper-and-pencil interviewing (PAPI) method to a computer-assisted interviewing (CAI) method, primarily to improve the quality of NHSDA estimates. This report has two purposes. First, it provides information on the impact of the redesign on the estimates produced from NHSDA. Researchers and other users of NHSDA data will find this information helpful in interpreting NHSDA estimates, particularly if they are interested in comparing data from the new design with data from the old design. The second purpose is to present research findings of interest to survey methodologists involved in designing and conducting surveys of all types, not just surveys of substance abuse.

METHODS: Implementation of these significant changes posed a number of difficult challenges involving a variety of methodological issues. These issues cover many aspects of the survey, including the management of fieldwork, the processing of data, and the reporting of results. This

publication discusses several of the most critical issues encountered and describes how the research team conducting the survey addressed them.

RESULTS/CONCLUSIONS: Although these findings taken as a whole could be considered a case study in the redesign of a major ongoing survey, several of the chapters in this report present important research findings that are applicable to many types of surveys.

Redesigning an ongoing national household survey: Methodological issues

CITATION: Gfroerer, J., Eyerman, J., & Chromy, J. (Eds.). (2002). *Redesigning an ongoing national household survey: Methodological issues* (DHHS Publication No. SMA 03-3768). Rockville, MD: Substance Abuse and Mental Health Services Administration, Office of Applied Studies.

PURPOSE/OVERVIEW: In 1999, a major redesign of the National Household Survey on Drug Abuse (NHSDA) was implemented involving both the sample design and the data collection method of the survey. The data collection method was changed from a paper-and-pencil interviewing (PAPI) method to a computer-assisted interviewing (CAI) method, primarily to improve the quality of NHSDA estimates. Implementation of these significant changes posed a number of difficult challenges involving a variety of methodological issues. These issues cover many aspects of the survey, including the management of fieldwork, the processing of data, and the reporting of results.

METHODS: This publication discusses several of the most critical issues encountered and describes how the research team conducting the survey addressed them. This report has two purposes. First, it provides information on the impact of the redesign on the estimates produced from NHSDA. Researchers and other users of NHSDA data will find this information helpful in interpreting NHSDA estimates, particularly if they are interested in comparing data from the new design with data from the old design. The second purpose is to present research findings of interest to survey methodologists involved in designing and conducting surveys of all types, not just surveys of substance abuse. Although these findings taken as a whole could be considered a case study in the redesign of a major ongoing survey, several of the chapters in this report present important research findings that are applicable to many types of surveys.

RESULTS/CONCLUSIONS: N/A.

Impact of interviewer experience on respondent reports of substance use

CITATION: Hughes, A., Chromy, J., Giacoletti, K., & Odom, D. (2002). Impact of interviewer experience on respondent reports of substance use. In J. Gfroerer, J. Eyerman, & J. Chromy (Eds.), *Redesigning an ongoing national household survey: Methodological issues* (DHHS Publication No. SMA 03-3768, pp. 161-184). Rockville, MD: Substance Abuse and Mental Health Services Administration, Office of Applied Studies.

PURPOSE/OVERVIEW: The redesign of the National Household Survey on Drug Abuse (NHSDA) in 1999 included a change in the primary mode of data collection from paper-and-

pencil interviewing (PAPI) to computer-assisted interviewing (CAI). In addition, the sample design changed from a Nation-based design to a State-based one, and a supplemental sample was collected using PAPI to measure change between 1999 and earlier years. The overall sample size increased from 25,500 in 1998 to 80,515 in 1999, including the CAI and supplemental PAPI samples. Consequently, it was necessary to hire more interviewers than in previous years, which resulted in a higher proportion of inexperienced interviewers. New interviewing staff turnover also was high in 1999, requiring additional training of newly hired interviewers and contributing to the general inexperience of the interviewing staff for both the CAI and PAPI samples.

METHODS: This chapter describes the analysis that was done to understand and explain the relationship between the effects of changes in mode, sample design, and interviewer staffing.

RESULTS/CONCLUSIONS: The analysis presented in this chapter indicates that the uneven mix of experienced and inexperienced NHSDA field interviewers (FIs) in 1999 had some effect on estimated substance use rates for that year. Overall, the effect on 1999 CAI prevalence estimates was smaller in magnitude than the effect on 1999 PAPI rates, which was an indication that the CAI methods played a role in reducing the effects of FI experience on substance use rates. However, because the mechanism of these effects was unknown, it was determined that additional studies would be undertaken to increase an understanding of this phenomenon. In the meantime, analyses of interviewer effect as seen in this chapter were to continue to be presented in subsequent reports. These findings resulted in an added emphasis—in training and in the field—on encouraging experienced and new FIs to follow the interview protocol.

Development of editing rules for CAI substance use data

CITATION: Kroutil, L., & Myers, L. (2002). Development of editing rules for CAI substance use data. In J. Gfroerer, J. Eyerman, & J. Chromy (Eds.), *Redesigning an ongoing national household survey: Methodological issues* (DHHS Publication No. SMA 03-3768, pp. 85-109). Rockville, MD: Substance Abuse and Mental Health Services Administration, Office of Applied Studies.

PURPOSE/OVERVIEW: A major change to the study protocol for the 1999 National Household Survey on Drug Abuse (NHSDA) was the shift from paper-and-pencil interviewing (PAPI) to computer-assisted interviewing (CAI). Although many of the substance use questions are similar in the two instruments, there are some differences. In addition, whereas the PAPI questionnaire required respondents to answer all questions in most sections, the CAI instrument makes extensive use of skip instructions. These significant differences in the nature of the data obtained in the new and old instruments necessitated the development of entirely new editing rules.

METHODS: This chapter discusses the development of the new editing rules for the NHSDA CAI data and presents the results of an investigation of alternative editing methods for CAI data. The analysis was based primarily on data from the first 6 months of data collection in 1999. The authors discuss data quality issues that were affected by conversion to CAI, present the general methodological approach used to define and test alternative rules for defining a usable case, discuss alternative editing rules, along with the final rule that was implemented, and present

selected substance use measures to compare the impact of the new CAI editing procedures with the former PAPI procedures.

RESULTS/CONCLUSIONS: The authors find for some substances that there was little change in going from the raw CAI data provided by the respondents to the final imputed estimates. For example, the estimate of marijuana use in the past month changed from 4.6 percent in the raw data to 4.7 percent after editing and imputation. In comparison, editing made a greater contribution to estimates of past month marijuana use and cigarette use in the 1998 data. For past month marijuana use, the raw estimate in 1998 was 4.0 percent (weighted), the estimate after editing was 5.0 percent, and the final imputed estimate was 5.0 percent. Thus, the editing procedures that had been used in the NHSDA since 1994 increased the 1998 estimate of past month marijuana use by about 25 percent relative to the raw data; the additional impact of imputation on the final estimate of past month marijuana use was virtually nil. Differences in the impacts of the 1999 CAI editing and imputation procedures and those used in prior years were even more pronounced for less commonly used substances. Following imputation, a total of 70 CAI respondents were classified as past month heroin users, or a net increase of only 4 cases relative to the raw and a net increase of 6 relative to the edited. In comparison, the editing procedures in 1998 nearly doubled the number of respondents classified as being past month heroin users (17 respondents in the raw data and an additional 11 cases who were assigned to this category through editing). In all, these changes in the 1999 CAI editing procedures represent an improvement over the way missing or inconsistent data had been handled in that these issues are resolved primarily through statistical methods.

2001 National Household Survey on Drug Abuse: Incentive experiment combined quarter 1 and quarter 2 analysis

CITATION: Office of Applied Studies. (2002, July). *2001 National Household Survey on Drug Abuse: Incentive experiment combined quarter 1 and quarter 2 analysis* (RTI/07190.388.100, prepared for the Office of Applied Studies, Substance Abuse and Mental Health Services Administration, by RTI under Contract No. 283-98-9008). Rockville, MD: Substance Abuse and Mental Health Services Administration.

PURPOSE/OVERVIEW: The purpose of this report was to summarize the results of the incentive experiment in the 2001 National Household Survey on Drug Abuse (NHSDA) and to evaluate the best treatment option for the use of monetary incentives in future NHSDAs. The NHSDA experienced a considerable decline in response rates in 1999 due in part to the transition from a national probability sample to a State probability sample designed to yield State-level estimates. A series of management adjustments were made to improve the response rates in 2000. In general, the adjustments were successful, and a recovery was made from the 1999 decline. However, the rates remained below the project target rate and the historical NHSDA average. An incentive payment to respondents was considered as an option for addressing the downward trend in respondent cooperation. However, it has been noted that incentive payments may have a negative impact on areas of data quality other than unit response rates (Shettle & Mooney, 1999). Although it may lead to better response rates, it is possible that the additional costs may exceed the constraints of the project budget. In an effort to understand the risks and benefits associated with a respondent incentive, the NHSDA's sponsor, the Substance Abuse and

Mental Health Services Administration (SAMHSA), requested that the NHSDA's contractor, Research Triangle Institute (RTI), conduct a special methodological field test in the form of an incentive experiment.

METHODS: The experiment was overlaid on the NHSDA main study data collection sample and scheduled during the first two quarters of 2001. A randomized, split-sample, experimental design was included with the main study data collection of the NHSDA to compare the impact of $20 and $40 incentive treatments with a $0 control group on measures of respondent cooperation, data quality, survey costs, and population substance use estimates. This report is the second of two. The first report describes the experimental design and the results from data collection in the first quarter of 2001 (Eyerman, Bowman, Odom, Vatalaro, & Chromy, 2001a). This second report provides combined findings for the full experiment for both quarters.

RESULTS/CONCLUSIONS: The results were very promising. The $20 and the $40 treatments produced significantly better interview response rates than the control for the combined results of both quarters of the experiment. This improvement led to a gain in overall response rates of about 10 points for each treatment. Furthermore, both the $20 and the $40 treatments more than paid for themselves, each resulting in a lower data collection cost per completed case, including incentive payment, than the control. The incentive payments had a favorable impact on measures of respondent cooperation. Both treatments had significantly lower refusal rates than the control's rate, and the $40 treatment had significantly lower noncontact rates than the control's. Field interviewers reported that the incentives reduced the amount of effort required to complete a case and that the payments influenced the respondent's decision to cooperate. Perhaps most importantly, the incentive payments had little impact on the population estimates of past month alcohol, cigarette, or marijuana use. The prevalence rates for past month use of these substances by respondents in the treatment groups were not significantly different from those reported by those in the control. This suggests that incentive payments encourage greater participation by respondents, but do not change their self-reported substance use. Incentives may thus improve estimates by reducing nonresponse bias without increasing response bias. Taken together, the results clearly favor a $40 incentive payment for all persons selected for the NHSDA.

Changes in NHSDA measures of substance use initiation

CITATION: Packer, L., Odom, D., Chromy, J., Davis, T., & Gfroerer, J. (2002). Changes in NHSDA measures of substance use initiation. In J. Gfroerer, J. Eyerman, & J. Chromy (Eds.), *Redesigning an ongoing national household survey: Methodological issues* (DHHS Publication No. SMA 03-3768, pp. 185-220). Rockville, MD: Substance Abuse and Mental Health Services Administration, Office of Applied Studies.

PURPOSE/OVERVIEW: The National Household Survey on Drug Abuse (NHSDA) data are used to generate a number of annual estimates relating to the initiation of substance use. Using the responses to retrospective questions about age of first use, annual estimates are generated for the incidence rate of first substance use, for the number of initiates to substance use, and for the average age of first use. Estimates of new initiates and average age of first use are reported for all lifetime substance users aged 12 or older. These use initiation measures are important because they can capture the rapidity with which new substance users arise in specific population

subgroups and can identify emerging patterns of substance use. The redesign of the NHSDA in 1999 introduced some changes in the questions about initiation, as well as the method of administration. In the presence of these changes, the overall data processing and estimation methodologies were reviewed and, in some cases, revised. (The revisions to the editing and imputation procedures are summarized in Chapters 5 and 6 of this 2002 report.) Limitations of the existing methodology for computing incidence rates were found. As a result, a new incidence rate methodology was developed. The definition of initiation of daily cigarette use was modified, and an adjustment to the program logic in the calculation for the incidence of first daily use of cigarettes also was made.

METHODS: This chapter is organized in three sections addressing the impact of methodological change on substance use initiation measures. Section 9.1 describes the old and new incidence rate estimation methods and evaluates its impact; the impact of the editing and imputing changes is evaluated in conjunction with the method impact. Section 9.2 focuses on the questionnaire wording and administration mode effects. Section 9.3 focuses on all the issues associated with initiation of first daily use of cigarettes.

RESULTS/CONCLUSIONS: Although the estimates for individual years were quite variable, the overall average impact of the new editing and imputation procedures was to increase incidence rates for both age groups (12 to 17 and 18 to 25) and to increase the estimated number of new initiates. The largest impacts were observed for pain relievers and other substances that use multiple gate questions before presenting the age-of-first-use question. Estimates of the average age of initiation did not appear to be consistently changed in either direction by the change in editing and imputation. The impact of the new method of incidence rate calculation also was studied. The number of new initiates occurring at age 17 was presumably quite high for almost all substances. The new incidence rate calculation rules treated respondents as 17 year olds right up to (but not including) their 18^{th} birthday. The old rule classified respondents as 18 years old for the entire year in which their 18^{th} birthday occurred. Thus, the new calculation method had the effect of increasing the estimates of time at risk and the number of initiates for 17 year olds, but because the number of initiates is high at age 17, the overall impact was greater on the numerator than the denominator. As a result, the incidence rates for persons aged 12 to 17 increased and the incidence rates for persons 18 to 25 usually decreased somewhat with the new method. Mode effects could not be cleanly isolated because of some accompanying changes in the question routing process and supplementary questions on date of first use for recent users that were implemented in conjunction with the implementation of computer-assisted interviewing (CAI). Within this limitation, comparable data from paper-and-pencil interviewing (PAPI) and CAI were studied. One somewhat surprising result was that the level of missing or inconsistent data actually increased with the introduction of CAI. However, this may have resulted because of the increased number of checks employed to identify inconsistent data in the post-survey processing. The increase in the proportion of missing age-at-first-use data may have been facilitated by the respondent's option to answer "Don't know" or "Refused" in CAI. A pattern of mode effects similar to that observed for reported lifetime substance use was found, with generally higher reporting of initiation in CAI than in PAPI.

Predictive mean neighborhood imputation for NHSDA substance use data

CITATION: Singh, A., Grau, E., & Folsom, R., Jr. (2002). Predictive mean neighborhood imputation for NHSDA substance use data. In J. Gfroerer, J. Eyerman, & J. Chromy (Eds.), *Redesigning an ongoing national household survey: Methodological issues* (DHHS Publication No. SMA 03-3768, pp. 111-133). Rockville, MD: Substance Abuse and Mental Health Services Administration, Office of Applied Studies.

PURPOSE/OVERVIEW: In 1999, the instrument used to administer the National Household Survey on Drug Abuse (NHSDA) was changed from a paper-and-pencil interviewing (PAPI) format to a computer-assisted interviewing (CAI) format. In previous years, imputation of missing values in recency of use and frequency of use in the past 12 months was accomplished with an unweighted sequential hot-deck procedure. In the spirit of improving the quality of estimates from the redesigned NHSDA and as a result of fundamental differences between PAPI and CAI, there was a need to change the way missing data were edited and imputed. The implementation of the "flag and impute" editing rule, described in this 2002 report's Chapter 5, and the desire to impute more variables required a new method that was rigorous, flexible, and preferably multivariate.

METHODS: This chapter presents a new imputation method with these characteristics, termed predictive mean neighborhoods (PMN), that was used to impute missing values in the NHSDA substance use variables. Following a discussion of background in Section 6.1, this chapter outlines the previously used hot-deck method, along with its limitations, in Section 6.2. The new method is described in general in Section 6.3, followed by details of the method in Section 6.4. Section 6.5 compares the method with other available methods and provides details concerning the motivation for employing a new method. In the concluding section (Section 6.6), the impact of imputation on substance use estimates is compared between PAPI and CAI.

RESULTS/CONCLUSIONS: The authors assess the relative impact of imputation for 1998 and 1999 (CAI) estimates of past month use, past year use, and lifetime use of all substances in the core section of the questionnaire. Because of numerous changes between the 1998 sample and the 1999 CAI sample, it would not be advisable to compare the final prevalence estimates (final percent) between the two samples. However, some comments can be made about the comparison of the "relative percent from imputes" between the two samples. In general, imputation had greater impact on the prevalence estimates in the CAI sample than on the estimates in the PAPI sample. With the implementation of the flag-and-impute editing rule in the CAI sample, where inconsistencies would be resolved by imputation, this result was not surprising. The exceptions to this rule were either due to differences in questionnaire format between PAPI and CAI or to attributes of the modules themselves.

1999–2001 National Household Survey on Drug Abuse: Changes in race and ethnicity questions

CITATION: Snodgrass, J. A., Grau, E. A., & Caspar, R. A. (2002, October 18). *1999–2001 National Household Survey on Drug Abuse: Changes in race and ethnicity questions* (RTI/07190.488.010, prepared for Office of Applied Studies, Substance Abuse and Mental Health Services Administration, Contract No. 283-98-9008). Research Triangle Park, NC: RTI.

PURPOSE/OVERVIEW: Since the inception of the National Household Survey on Drug Abuse (NHSDA), renamed the National Survey on Drug Use and Health (NSDUH) as of 2002, questions have been included to determine the race and ethnicity of each respondent. Race and ethnicity are routinely used as part of the demographic breakdowns in the analyses and the various reports generated from the survey. From 1971 to 1998, the race and ethnicity questions underwent few changes. (See Appendix A of this 2002 report for the full list of race and ethnicity questions used for each NHSDA survey year from 1971 to 2001.) However, along with the switch from paper-and-pencil interviewing (PAPI) methods of questionnaire administration to computer-assisted interviewing (CAI) methods in 1999, the race and ethnicity categories were updated pursuant to new Office of Management and Budget (OMB) directives.

METHODS: This report details the revisions to the race and ethnicity questions. The report includes the history of the change, how the change affected the editing and imputation procedures, and how it changed the derivation of the race and ethnicity variables used in NHSDA analyses.

RESULTS/CONCLUSIONS: N/A.

Substance abuse among older adults in 2020: Projections using the life table approach and the National Household Survey on Drug Abuse

CITATION: Woodward, A. (2002, December). Substance abuse among older adults in 2020: Projections using the life table approach and the National Household Survey on Drug Abuse. In S. P. Korper & C. L. Council (Eds.), *Substance use by older adults: Estimates of future impact on the treatment system* (DHHS Publication No. SMA 03-3763, Analytic Series A-21, pp. 95-105). Rockville, MD: Substance Abuse and Mental Health Services Administration, Office of Applied Studies.

PURPOSE/OVERVIEW: One way of projecting substance abuse problems among older adults is to use a life table approach. The National Household Survey on Drug Abuse (NHSDA), a major data source on substance use and abuse among the U.S. civilian population aged 12 or older, could potentially be used in a life table approach.

METHODS: In the life table approach, cohorts are followed for a given period of time to determine their various outcomes. In the table, one cohort's drug use is followed as the cohort ages. The NHSDA data are reviewed to see whether they can be used in this fashion to produce estimates for the groups who used illicit drugs or drank heavily or who were substance dependent.

RESULTS/CONCLUSION: A review of the NHSDA shows that, even with its increase in sample size in 1999, the survey currently does not provide sufficient detailed data to be used in a life table approach. The survey could be expanded, however, with selected questions added in a special supplement so that a life table or other more sophisticated approach could be used to make projections of substance abuse problems among older adults.

Summary of NHSDA design changes in 1999

CITATION: Wright, D., Barker, P., Gfroerer, J., & Piper, L. (2002). Summary of NHSDA design changes in 1999. In J. Gfroerer, J. Eyerman, & J. Chromy (Eds.), *Redesigning an ongoing national household survey: Methodological issues* (DHHS Publication No. SMA 03-3768, pp. 9-22). Rockville, MD: Substance Abuse and Mental Health Services Administration, Office of Applied Studies.

PURPOSE/OVERVIEW: An entirely new sample design and a state-of-the-art data collection methodology were implemented with the 1999 National Household Survey on Drug Abuse (NHSDA). The sample design changed from a national, stratified, multistage area probability sample to a 50-State design, with independent stratified, multistage area probability samples selected in each State. The sample size increased from about 25,500 interviews in 1998 to about 67,000 interviews in 1999. For the first time in NHSDA history, the 1999 survey administered the interview using computer-assisted interviewing (CAI) technology exclusively, including both computer-assisted personal interviewing (CAPI) and audio computer-assisted self-interviewing (ACASI). Because this new methodology was being implemented, an additional national sample was selected, and about 14,000 interviews administered using the previous paper-and-pencil interviewing (PAPI) methodology.

METHODS: Together, the PAPI sample and the CAI sample served three purposes. First, the PAPI samples for 1998 and 1999 provided a way to continue to measure the trend in substance use for that period. Second, with both representative samples for 1999, the effect of the change in data collection from PAPI to CAI could be measured without being confounded with the measurement of trend. Third, with a measurement of the impact of the switch to the CAI methodology, estimates for 1998 and earlier years could be adjusted to be comparable with CAI estimates for 1999 and later so that long-term trends in substance use could be estimated.

The CAI and PAPI samples for 1999 together resulted in 81,000 completed interviews. The 1999 NHSDA fully employed another technological innovation: use of a hand-held computer at each sample dwelling unit to conduct household screening and to select the sample person(s) for the interview. With this new design, technology, and markedly increased sample size, the structure of the data collection staff also had to be modified significantly for 1999. This chapter presents details of these changes.

RESULTS/CONCLUSIONS: N/A.

2003

Possible age-associated bias in reporting of clinical features of drug dependence: Epidemiological evidence on adolescent-onset marijuana use

CITATION: Chen, C. Y., & Anthony, J. C. (2003). Possible age-associated bias in reporting of clinical features of drug dependence: Epidemiological evidence on adolescent-onset marijuana use. *Addiction, 98*(1), 71-82.

PURPOSE/OVERVIEW: As of 2003, the latest drug use reports showed a higher level of marijuana dependence for adolescents than for adults. This paper explored potential age-related differences in marijuana dependence using multivariate analysis and item-response biases.

METHODS: Marijuana dependence was measured using seven binary survey items from the 1995 to 1998 National Household Survey on Drug Abuse (NHSDA) questionnaires. Of the 86,021 respondents for these 4 years of the NHSDA, 2,628 (1,866 adolescents and 762 adults) were identified as recent-onset marijuana users. Multivariate response analysis using a generalized linear model (GLM) and generalized estimating equations (GEE) was performed to measure the age-related difference in reported marijuana use while taking into account the interdependencies of the yes-no responses and controlling for covariates. Further analyses were performed using the multiple indicators/multiple causes (MIMIC) multivariate response model to measure age-associated response bias.

RESULTS/CONCLUSIONS: The primary findings were that of the recent-onset marijuana users, younger users reported more drug dependency than older users. This association was found to be statistically significant in the multivariate analysis model, which controlled for other covariates. The MIMIC model found that there also were age-related biases in reporting of drug use. Younger users were biased toward reporting dependent behaviors. These findings support the notion that there may be differences in what constitutes "marijuana dependence" for adolescents compared with adults.

The utility of debriefing questions in a household survey on drug abuse

CITATION: Fendrich, M., Wislar, J. S., & Johnson, T. P. (2003). The utility of debriefing questions in a household survey on drug abuse. *Journal of Drug Issues, 33*(2), 267-284.

PURPOSE/OVERVIEW: Respondent's discomfort in revealing answers to sensitive survey questions is often a cause of underreporting. This paper measures the effect that respondents' comfort level has on their answers to sensitive questions and whether there are any differences in mode.

METHODS: Respondents were randomly assigned to receive either the control or experimental condition. The control condition replicated the design procedures from the 1995 National Household Survey on Drug Abuse (NHSDA), which used an interviewer-administered paper-

and-pencil interviewing (PAPI) survey. The experimental condition had the choice of completing a computer-assisted personal interviewing (CAPI) survey or an audio computer-assisted self-interviewing (ACASI) survey. Following the interview, respondents were asked for a hair sample. At the end of the interview, interviewers administered two types of debriefing questions, subjective and projective, to elicit respondents' level of discomfort during the study.

RESULTS/CONCLUSIONS: Respondents' willingness to disclose sensitive information was correlated with their subjective and projective levels of discomfort with the survey. Contrary to the hypothesis, respondents who disclosed information revealed less comfort with the survey on the subjective debriefing items and more comfort on the projective debriefing questions. In addition, the degree of discomfort for survey respondents was greater for those in the experimental condition than the control condition. Levels of projective and subjective discomfort differed by subgroups for race, age, and education.

Substance use treatment need among older adults in 2020: The impact of the aging baby-boom cohort

CITATION: Gfroerer, J., Penne, M., Pemberton, M., & Folsom, R. (2003). Substance use treatment need among older adults in 2020: The impact of the aging baby-boom cohort. *Drug and Alcohol Dependence, 69,* 127-135.

PURPOSE/OVERVIEW: The purpose of this paper is to provide estimates of the number of older adults (defined as those 50 or older) needing treatment for substance use problems in the future as the U.S. baby-boom population ages.

METHODS: Data from the 2000 and 2001 National Household Surveys on Drug Abuse (NHSDAs) were used to estimate logistic regression models among persons aged 50 or older, predicting treatment need in the past 12 months, which was defined as being classified with substance dependence or abuse based on DSM-IV criteria. Separate regression models were estimated for persons depending on whether they had used alcohol by age 30. For those who had not used alcohol by age 30, age was the only predictor. For those who had used alcohol prior to age 31, predictors in the regressions included age, gender, race/ethnicity, and substance use prior to age 31. Estimated parameters from these models then were applied to a pooled sample of persons 30 or older from the 2000 NHSDA and 31 or older from the 2001 NHSDA to produce predicted probabilities of treatment need. Weights were adjusted to match census population projections on age, race, and gender. An additional adjustment was made for expected mortality rates. Standard errors for the 2020 projections were computed using jackknife replication methods.

RESULTS/CONCLUSIONS: The estimated number of older adults needing treatment for a substance abuse problem is projected to increase from 1.7 million in 2000/2001 to 4.4 million in 2020. This increase is the result of a 50 percent increase in the population aged 50 or older (from 74.8 million to 112.5 million) combined with a 70 percent increase in the rate of treatment need (from 2.3 to 3.9 percent) in the older population. Increases are projected for all gender, race, and age groups. About half of the projected 2020 population needing treatment will be aged 50 to 59, and two thirds will be male.

Comparison of linear and nonlinear models for generalized variance functions for the National Survey on Drug Use and Health

CITATION: Gordek, H., & Singh, A. C. (2003). Comparison of linear and nonlinear models for generalized variance functions for the National Survey on Drug Use and Health. In *Proceedings of the 2003 Joint Statistical Meetings, American Statistical Association, Section on Survey Research Methods, San Francisco, CA* [CD-ROM]. Alexandria, VA: American Statistical Association.

PURPOSE/OVERVIEW: The generalized variance function (GVF) provides an approximation to the variance of an arbitrary domain estimate by means of modeling the relationship between the estimated variances of a selected set of domain estimates and a set of covariates, rather than through direct computation. The use of GVF is important in surveys with a very large number of characteristics, where publishing limitations and the immense amount of potential subgroupings may make several variance estimates of interest unavailable directly.

METHODS: For the National Survey on Drug Use and Health (NSDUH), GVFs were obtained using linear models with the log of the relative variance as the dependent variable, as well as a somewhat modified version that ensures the resulting design effect (deff) to be at least one. The authors consider a nonlinear generalization of these models. Numerical results on comparison of various models using the 2001 NSDUH data are presented.

RESULTS/CONCLUSIONS: The authors extended the idea of constant deff for a subset of estimates to constant deff-type parameters, such as variance-odds deff, and then proposed an alternative GVF model that overcomes the limitations of existing models. Ordinary least squares was used to fit these models to the NSDUH data, and results from different models were compared. It was found that different models performed quite similarly, and there seemed no compelling reason to change from simpler, commonly used models. However, the authors concluded that it would be useful to investigate the impact of using other models, as well as the use of weighted least squares in model fitting. For instance, the authors indicated that the effect of nonlinear modeling should be investigated. More specifically, rather than formulating a linear model for the transformed point estimate of variance, one can model linearly the mean of the variance estimate after transformation (such as log or logit).

Screening for serious mental illness in the general population

CITATION: Kessler, R. C., Barker, P. R., Colpe, L. J., Epstein, J. F., Gfroerer, J. C., Hiripi, E., Howes, M. J., Normand, S. L., Manderscheid, R. W., Walters, E. E., & Zaslavsky, A. M. (2003). Screening for serious mental illness in the general population. *Archives of General Psychiatry, 60*(2), 184-189.

PURPOSE/OVERVIEW: The Substance Abuse and Mental Health Services Administration (SAMHSA) was charged with the task of defining "serious mental illness" (SMI) in adults and establishing a method that States can use to estimate its prevalence.

METHODS: This paper compared three sets of screening scales used to estimate the prevalence of SMI: the World Health Organization (WHO) Composite International Diagnostic Interview Short-Form (CIDI-SF) scale, the K10 and K6 psychological distress scales, and the WHO Disability Assessment Schedule (WHO-DAS). A convenience sample of 155 respondents received the three screening scales using computer-assisted self-interviewing (CASI). After the self-administered questions, respondents completed the 12-month Structured Clinical Interview (SCID), including the Global Assessment of Functioning (GAF). SMI was defined in the SCID as any 12-month DSM-IV disorder (defined in the 4th edition of the *Diagnostic and Statistical Manual of Mental Disorders*), besides substance use, that also had a GAF score less than 60.

RESULTS/CONCLUSIONS: All screening scales were found to be significantly correlated with SMI. The shortest scale, the K6, was the most statistically significant predictor of SMI. These results support the notion that short, fully developed, and carefully constructed screening scales can be strong predictors of the same results found in more lengthy and expensive clinical interviews. Another advantage of the K6 and K10 scales is that they can be administered in less time than 2 or 3 minutes, respectively.

The effect of interviewer experience on the interview process in the National Survey on Drug Use and Health

CITATION: Odom, D. M., Eyerman, J., Chromy, J. R., McNeeley, M. E., & Hughes, A. L. (2003). The effect of interviewer experience on the interview process in the National Survey on Drug Use and Health. In *Proceedings of the 2003 Joint Statistical Meetings, American Statistical Association, Section on Survey Research Methods, San Francisco, CA* [CD-ROM]. Alexandria, VA: American Statistical Association.

PURPOSE/OVERVIEW: Analysis of survey data from the National Survey on Drug Use and Health (NSDUH) has shown a relationship between field interviewer (FI) experience, response rates, and the prevalence of self-reported substance use (Eyerman, Odom, Wu, & Butler, 2002; Hughes, Chromy, Giacoletti, & Odom, 2001, 2002; Substance Abuse and Mental Health Services Administration [SAMHSA], 2000). These analyses have shown a significant and positive relationship between the amount of prior experience an FI has with collecting NSDUH data and the response rates that a FI produces with his or her workload. These analyses also have shown a significant and negative relationship between the amount of prior experience of an FI and the prevalence of substance use reported in cases completed by that FI. In general, these analyses have been consistent with the published literature that FIs can influence both the success of the data collection process and accuracy of the population estimates (Martin & Beerteen, 1999; Singer, Frankel, & Glassman, 1983; Stevens & Bailar, 1976). Previous NSDUH analyses examined response rates and prevalence estimates independently. This made it difficult to determine whether the lower prevalence estimates for experienced FIs were a result of the change in the sample composition due to higher response rates or whether the lower prevalence estimates were a result of a direct effect of FI behavior on respondent self-reporting.

METHODS: This analysis combines these two explanations to produce a conceptual model that summarizes the authors' expectations for the relationship between FI experience and prevalence estimates. The combined explanation from the conceptual model is evaluated in a series of

conditional models to examine the indirect effect of response rates and the direct effect of FI experience on prevalence estimates. Earlier analyses using NSDUH data have shown a negative correlation between interviewer experience and substance use rates. However, these studies only examined the last step in the interviewing process, administering the questionnaire. This paper further explores the effect of interviewer experience by investigating a series of separate logistic models that are conditionally based on each step of the screening and interviewing (S&I) process. The S&I steps examined are contacting the household, gaining household cooperation, contacting the selected person(s), interviewing the selected person(s), and reporting of substance use. By separating the analysis into these steps, estimating the effect of interviewer experience on data collection at each stage of the survey is possible.

RESULTS/CONCLUSIONS: In conclusion, the analysis shows that increased FI experience simultaneously increases response rates and decreases prevalence estimates. In addition, the effect of increased FI experience on prevalence estimates cannot be fully explained by the adjustments based on earlier models (i.e., S&I level) to the final prevalence estimate model. In other words, the FI effect on prevalence cannot be fully attributed to the increase in response rates by experienced FIs. Furthermore, FI experience was significant in the final model, showing that the covariates also did not account for all the decrease in prevalence. Three hypotheses were given as possible explanations for the decrease in prevalence. As was shown in the statistical analysis, the marginal rates were too extreme to support the first hypothesis. This means that although some level of selection bias may be occurring, it is not the only cause of the decrease in prevalence estimates for experienced FIs. More likely, the relationship between FI experience and prevalence estimates is captured in hypothesis 3 (i.e., the decrease in prevalence estimates for experienced FIs is a function of lower substance use reporting by the additional respondents they obtain and also the remaining respondents that FIs with all levels of experience interview).

Appendix C: NSDUH changes and their impact on trend measurement

CITATION: Office of Applied Studies. (2003). Appendix C: NSDUH changes and their impact on trend measurement. In *Results from the 2002 National Survey on Drug Use and Health: National findings* (DHHS Publication No. SMA 03-3836, NSDUH Series H-22, pp. 107-137). Rockville, MD: Substance Abuse and Mental Health Services Administration.

PURPOSE/OVERVIEW: This appendix presents the results of analyses designed to determine the degree to which increases in the rates of substance use, dependence, abuse, and serious mental illness (SMI) between the 2001 National Household Survey on Drug Abuse (NHSDA) and the 2002 National Survey on Drug Use and Health (NSDUH) could be attributed to important methodological differences between the two surveys. Major changes between the 2 years included (1) a change in the survey's name, (2) providing a $30 incentive to all interview respondents, (3) improved data collection quality control procedures, and (4) switching to the use of 2000 census data as the basis for population weighting adjustments for the 2002 survey.

METHODS: Six types of analyses were used to assess the degree to which methodological changes could account for the various increases in prevalence rates: (1) a retrospective cohort analysis that examined the degree to which increases in lifetime use could be attributed to new initiates; (2) a response rate pattern analysis that examined changes in response rates between the

fourth quarter of 2001 and the first quarter of 2002 for various geographic and demographic groups, reasons for refusal, and field interviewer characteristics; (3) a response rate impact analysis that attempted to determine whether the increased response rate between 2001 and 2002 resulted in "additional" respondents who in turn accounted for higher rates of substance use; (4) an analysis of the impact of new census data in which data for the 2001 survey were reweighted using 2000-based census control totals rather than 1990-based census data; (5) analyses in which measures of substance use were regressed on variables that included indicators related to the timing and occurrence of changes in data collection quality control procedures to determine the impact of these changes; and (6) a reexamination of differences in prevalence rates from the 2001 incentive experiment by applying response propensity adjusted weights to account for response rate differences and a comparison of these differences with differences in estimates between quarter 4 of the 2001 survey and quarter 1 of the 2002 survey.

RESULTS/CONCLUSIONS: (1) Increases in lifetime use cannot be accounted for by new initiates in 2002 or a cohort shift. (2) Response rate increases occurred across all geographic and demographic groups, with the exception of persons aged 50 or older. (3) "Additional" respondents from the response rate increase between 2001 and 2002 cannot account for the increase in prevalence rates between 2001 and 2002. (4) Reweighting the 2001 survey by using 2000 census-based control totals had minor effects on prevalence rates and a larger impact on totals. (5) Field interventions introduced during the 2001 survey appear to have had little effect on substance use estimates. (6) Differences in estimates between quarter 1 of 2002 and quarter 4 of 2001 were generally larger than differences in estimates from the incentive experiment between incentive and no-incentive groups, suggesting that incentive effects alone do not account for overall differences.

2004

Triad sampling in household surveys

CITATION: Aldworth, J., & Chromy, J. (2004). Triad sampling in household surveys. In *Proceedings of the 2004 Joint Statistical Meetings, American Statistical Association, Section on Survey Research Methods, Toronto, Canada* [CD-ROM]. Alexandria, VA: American Statistical Association.

PURPOSE/OVERVIEW: The motivation for selecting three persons from a dwelling unit (DU) derives from an interest in studying the behavioral relationships among certain triads (e.g., two parents and a child, a parent and two children). Computer-assisted DU screening provides the mechanism for targeted sampling of individuals, pairs, and even triads within a DU. Chromy and Penne (2002) showed how a modification of Brewer's (1963, 1974) method for samples of size two was used to select samples of 0, 1, or 2 persons from eligible DUs. They also developed a second adaptation to control the number of pairs selected.

METHODS: Chromy and Penne's modification of Brewer's method and their adaptation to control the number of persons selected from a DU is extended to the case of sampling triads within DUs. Some empirical data on household roster composition, and response rates based on the number of persons selected, are presented from the National Survey on Drug Use and Health (NSDUH).

RESULTS/CONCLUSIONS: The results of simulations of the sample selection and response process are presented as a means of evaluating alternatives. Possible negative impacts of triad selections on design effect, response rates, and response bias were assessed. The results indicate that although pair selections in the 2002 NSDUH show little negative impact, a field test assessment is needed for triad assessment.

Estimating substance abuse treatment need by state [editorial]

CITATION: Gfroerer, J., Epstein, J., & Wright, D. (2004). Estimating substance abuse treatment need by state [editorial]. *Addiction, 99*(8), 938-939.

PURPOSE/OVERVIEW: This editorial comments on the methods used by the Substance Abuse and Mental Health Services Administration (SAMHSA) to measure the need for the treatment of substance abuse disorders at the State level.

METHODS: The two methods compared are the one used in the National Survey on Drug Use and Health (NSDUH) and the index method proposed by W. E. McAuliffe and R. Dunn in this issue of *Addiction*. The authors describe the strengths and weaknesses of each of these methods.

RESULTS/CONCLUSIONS: The strengths of NSDUH are that it uses an independent probability sample for each State, data collection procedures and methods are defined and carried out consistently across States, and the data are current, reflecting well-defined time periods. A

validation study also has demonstrated that NSDUH has small biases and allows for estimates of different subgroups. The main limitations are the small sample sizes in each State and possible underreporting and undercoverage. The primary weakness of the index method is that their measures are not well defined and therefore not consistently interpreted. For these reasons, the NSDUH is recognized as more valid than the index method, but the NSDUH results are meant to be interpreted along with the results from other data sources.

Estimating trends in substance use based on reports of prior use in a cross-sectional survey

CITATION: Gfroerer, J., Hughes, A., Chromy, J., Heller, D., & Packer, L. (2004). Estimating trends in substance use based on reports of prior use in a cross-sectional survey. In S. B. Cohen & J. M. Lepkowski (Eds.), *Eighth Conference on Health Survey Research Methods* (DHHS Publication No. PHS 04-1013, pp. 29-34). Hyattsville, MD: U.S. Department of Health and Human Services, Public Health Service, Centers for Disease Control and Prevention, National Center for Health Statistics.

PURPOSE/OVERVIEW: Substance use trends in the United States have shown dramatic shifts since the 1960s. Among youths aged 12 to 17, the rate of past month marijuana use was less than 2 percent in the early 1960s, increased to 14 percent by 1979, then decreased to 3.4 percent in 1992 before rising to 8.2 percent in 1995. Major shifts in prevalence at different points in time and for different age groups have been observed for other substances, including cocaine, LSD, Ecstasy, opiates, cigars, and cigarettes (Substance Abuse and Mental Health Services Administration [SAMHSA], 2003). Accurate measurement of these trends is critical for policymakers targeting limited resources efficiently toward emerging problems. Trend data also are used for assessing the impact of prevention and treatment programs. The typical method used for measuring substance use trends is comparing prevalence estimates across repeated cross-sectional surveys. An alternative approach is to collect data about prior substance use within a cross-sectional survey and to construct prevalence estimates for prior years based on these data. Besides the cost advantages, these retrospective estimates have some analytic advantages. When data are obtained for different periods from the same respondents, trend analyses are more powerful, due to the positive correlation between estimates, as is the case in a longitudinal study. Retrospective estimates also may be the only alternative if estimates are needed for periods for which direct estimates are not available. However, retrospective estimates do have important limitations. Bias due to recall decay, telescoping, and reluctance to admit socially undesirable behaviors could cause underestimation or distort trends (Johnson, Gerstein, & Rasinski, 1998; Kenkel, Lillard, & Mathios, 2003). Bias also could result from coverage errors affecting the capability of the sample to represent the population of interest for prior time periods, due to mortality, immigration, or other changes in the population. This paper discusses several types of retrospective estimates and presents analyses of data from the National Survey on Drug Use and Health (NSDUH) to assess biases in these estimates.

METHODS: Several analyses were undertaken to assess bias in retrospective estimates. One known source of bias in retrospective estimates is the inclusion of data from immigrants who were not living in the United States in some prior years. The authors compared estimates of incidence and lifetime use (for those aged 12 to 17 or 18 to 25) for the full 2002 NSDUH sample

with estimates based on the sample excluding these immigrants, according to questions on country of birth and years in the United States.

Trends in incidence estimates for 1965 to 1990 based on 1991 to 1993 data (shortest recall), 1994 to 1998 data, 1999 to 2001 data, and 2002 data (longest recall) were compared. For the 1991 to 1997 period, trends based on the 1994 to 1998 data, 1999 to 2001 data, and 2002 data were compared. Consistency was assessed through visual inspection of curves and with correlations. Because of methodology changes, comparisons of levels from different sets of surveys were not made.

The authors compared 2002-based retrospective lifetime use estimates (excluding immigrants) with direct lifetime use estimates from earlier NSDUHs (for those aged 12 to 17 or 18 to 25) and from the Monitoring the Future (MTF) study, a survey of high school seniors (Johnston, O'Malley, & Bachman, 2003). To reduce the effect of sampling error, the authors combined several years of data, depending on availability, and generated average annual lifetime prevalence for specific time periods. Because 1999 and 2002 survey changes resulted in increased reporting of lifetime use, the authors expected retrospective estimates to be greater than the direct estimates for years before 1999.

The authors compared retrospective lifetime use estimates for 2002, based on 2003 NSDUH data (first 6 months of data currently available), to direct 2002 lifetime use estimates, from the 2002 NSDUH (first 6 months of data, for consistency). Comparisons were made for 19 substances for those aged 12 to 17 or 18 to 25.

To assess the accuracy of these estimates, the authors compared January to June 2003-based retrospective estimates of past year use in January to June 2002 to direct past year estimates from the January to June 2002 data, by age group.

RESULTS/CONCLUSIONS: Marijuana incidence estimates for 1965 to 2001 were 2.5 percent higher when immigrants were included. For most other illicit drugs, the bias was smaller, indicating that very little initiation for these drugs occurs among immigrants prior to their entry to the United States. However, biases for alcohol and cigarette incidence estimates were larger (8 percent for alcohol, 7 percent for cigarettes). In general, they were largest for the years 1979 to 1994 (3.5 percent for marijuana, 11 percent for alcohol, 10 percent for cigarettes) and smallest for years after 1997 (1 percent for marijuana, 3 percent for alcohol, and 3 percent for cigarettes). For lifetime prevalence rates, bias due to including immigrants was negative for nearly every substance because of the low rates of substance use among immigrants. For youth estimates during the period from 1979 to 1990, the inclusion of immigrants resulted in biases of about –14 percent for marijuana, –15 percent for cocaine, –9 percent for cigarettes, and –9 percent for alcohol. Bias was generally worse for estimates for those aged 12 to 17 than for those aged 18 to 25, and there was very little bias in any estimates for years after 1997. Estimates of alcohol use for those aged 18 to 25 including immigrants showed very small but positive bias (1.5 percent) for the period from 1982 to 1993 and a larger positive bias (7 percent) for 1965 to 1981.

A simple evaluation of the imputation procedures used in NSDUH

CITATION: Grau, E., Frechtel, P., Odom, D., & Painter, D. (2004). A simple evaluation of the imputation procedures used in NSDUH. In *Proceedings of the 2004 Joint Statistical Meetings, American Statistical Association, Section on Survey Research Methods, Toronto, Ontario, Canada* [CD-ROM]. Alexandria, VA: American Statistical Association.

PURPOSE/OVERVIEW: The National Survey on Drug Use and Health (NSDUH) is the primary source of information on drug use in the United States. Since 1999, the predictive mean neighborhoods (PMN) procedure has been used to impute missing values for many of the analytical variables. This method is a combination of two commonly used imputation methods: a nearest neighbor hot-deck and a modification of Rubin's predictive mean matching method. Although PMN has many practical advantages, it has not been evaluated formally.

METHODS: The authors proposed a simple simulation to evaluate PMN. Using only complete data cases, they induced random patterns of missingness in the data for selected outcome variables. Imputations then were conducted using PMN and a weighted nearest-neighbor hot-deck. This process of inducing missingness and imputing missing values was repeated multiple times. The imputed values using PMN and the weighted hot deck then were compared with the true values found in the complete data across the repeated iterations. In particular, the authors compared the number of matches between the two methods, as well as compared statistics derived from the data, such as drug prevalence estimates.

RESULTS/CONCLUSIONS: This evaluation showed that using PMN provided a modest advantage in the match test, and, in the missing completely at random (MCAR) case, using weighted sequential hot deck (WSHD) provided a modest advantage in the mean test. In each of these cases, however, the degree of difference between the methods, though significant, was not substantial. Not surprisingly, all methods did badly in the mean test when the data were not missing at random (NMAR).

Measuring treatment needs: A reply to Gfroerer, Epstein, and Wright

CITATION: McAuliffe, W. E. (2004). Measuring treatment needs: A reply to Gfroerer, Epstein, and Wright. *Addiction, 99*(9), 1219-1220.

PURPOSE/OVERVIEW: This commentary is a reply to the editorial titled "Estimating Substance Abuse Treatment Need by State" by Joe Gfroerer, Joan Epstein, and Doug Wright. The author commented on the strengths and weaknesses of studies used to measure the treatment need for substance abuse disorders.

METHODS: The two studies compared were the National Survey on Drug Use and Health (NSDUH) and the Substance Abuse Need Index (SNI) produced by W. E. McAuliffe and R. Dunn (2004). The author discussed the strengths and weaknesses of each approach.

RESULTS/CONCLUSIONS: NSDUH suffers from underreporting, undercoverage, and nonresponse, which likely leads to undercounting users of hard-core drugs, such as cocaine. The

SNI more accurately captures the incarcerated and homeless populations, improving their estimates for cocaine and drug abuse. However, the NSDUH estimates probably are more accurate for marijuana use. The author disagreed with Gfroerer et al. that the SNI was less reliable or valid than NSDUH and expressed the opinion that multiple data sources need to be considered together to get the full picture on drug abuse and need for treatment.

Substance abuse treatment needs and access in the USA: Interstate variations

CITATION: McAuliffe, W. E., & Dunn, R. (2004). Substance abuse treatment needs and access in the USA: Interstate variations. *Addiction, 99*(8), 999-1014.

PURPOSE/OVERVIEW: This paper analyzed two measures for substance abuse treatment needs across States. They were used to explore geographic variations in substance abuse, the causes for substance abuse, the stability of these estimates over time, and whether the severity of substance abuse was correlated to need.

METHODS: This study used alcohol dependency estimates from the National Survey on Drug Use and Health (NSDUH), and drug and alcohol use indexes based on alcohol-related mortality and arrests data, to measure interstate differences between substance abuse treatment needs and treatment services. This study tested the reliability and the validity of the survey measures used. The index for substance abuse treatment then was regressed on the measures for substance abuse need to identify differences in availability of treatment across States.

RESULTS/CONCLUSIONS: The individual indicators of treatment needs and availability of treatment for substance abuse across the United States had reliability and construct validity. Substance abuse problems in the United States are clustered geographically, with the most severe problems appearing in the western States. The biggest gaps between treatment need and treatment access appeared in the South where there was moderate need for treatment, but very low access. The interstate discrepancies in treatment need versus treatment services indicated that substance abuse problems in rural areas are often overlooked by treatment services.

Prevalence of adult binge drinking: A comparison of two national surveys

CITATION: Miller, J. W., Gfroerer, J. C., Brewer, R. D., Naimi, T. S., Mokdad, A., & Giles, W. H. (2004). Prevalence of adult binge drinking: A comparison of two national surveys. *American Journal of Preventive Medicine, 27*(3), 197-204.

PURPOSE/OVERVIEW: This paper compares two national surveys—the Behavioral Risk Factor Surveillance System (BRFSS) and the National Survey on Drug Use and Health (NSDUH)—that measure the prevalence of binge drinking across States and overall for the United States. The authors examine methodological differences between the two studies to assess their impact on the survey estimates.

METHODS: The main methodological difference between the two studies is that BRFSS is conducted over the telephone and NSDUH is conducted in-person using audio computer-assisted self-interviewing (ACASI). BRFSS assesses binge drinking using three alcohol questions, and

NSDUH uses nine; however, the questions are very similar. Because BRFSS uses a telephone sample and NSDUH is an in-person survey, data from NSDUH were limited to only telephone-households to make comparisons consistent. In addition, NSDUH was restricted to respondents aged 18 years or older to match BRFSS. Response rates, survey size, and other characteristics of the two studies were compared.

RESULTS/CONCLUSIONS: BRFSS binge drinking estimates for the United States and most States were considerably lower than the NSDUH estimates, even after stratifying for sociodemographic variables. The demographic characteristics of the sample were very similar; the majority were male and white, non-Hispanic. However, there were no significant differences for binge drinking in the past 30 days between the two studies. The large differences in estimates was likely due to the use of ACASI in NSDUH, which is perceived as more anonymous and yields higher reporting of sensitive behaviors. Other possible explanations for the differences are the sample size, the number of questions asked about alcohol, and the overall topic of the survey.

A system for detecting interviewer falsification

CITATION: Murphy, J., Baxter, R. K., Eyerman, J., Cunningham, D., & Barker, P. (2004). A system for detecting interviewer falsification. In *Proceedings of the 2004 Joint Statistical Meetings, American Statistical Association, Section on Survey Research Methods, Toronto, Ontario, Canada* [CD-ROM]. Alexandria, VA: American Statistical Association.

PURPOSE/OVERVIEW: The National Survey on Drug Use and Health (NSDUH) has developed a detailed system for the detection of interviewer falsification. This process includes phone, mail, and in-person field verification procedures, as well as the review of interview and interviewer-level process data to identify cases and interviewers requiring extra verification efforts. Although these components of the system successfully identify a majority of potential falsifiers, more savvy falsifiers may be able to remain undetected if they are aware that their process data are being scrutinized. To address this gap, NSDUH added a new component to the falsification detection system: the regular review of interview response and question-level timing data. This paper details the structure and operationalization of this system and presents examples of its effectiveness on NSDUH.

METHODS: Based on what is known about the area in which an interviewer is working and the types of cases he or she is assigned, a likely range of interview responses is calculated. Response distributions are compared with this likely range at the interviewer level to identify interviewers whose responses appear to be highly unlikely, given their caseloads. These additional measures make it even more difficult for falsifiers to remain undetected because they would need to have a specific understanding of the prevalence and correlates of substance use in order to enter likely responses. Similarly, question-level timings for particular items that require certain interviewer-respondent interactions are compared with a "gold standard" to detect outliers. Once potential falsifiers are identified, the work of the suspected interviewers is subject to 100 percent and/or in-person verification.

RESULTS/CONCLUSIONS: The analyses in this paper showed that falsification detection can be improved through the systematic review of response data and metadata, such as module and

item timings. Through early detection and remediation, the threat of falsification to survey bias and increased costs can be reduced. Although the NSDUH system has improved based on these recent enhancements, there are still many more enhancements that could be incorporated. In particular, the authors planned to assess the following techniques for possible adoption: (1) analysis of screening data, (2) analysis of record of calls data and other metadata, (3) a data mining approach, and (4) statistical process control.

Nonresponse among persons age 50 and older in the National Survey on Drug Use and Health

CITATION: Murphy, J., Eyerman, J., & Kennet, J. (2004). Nonresponse among persons age 50 and older in the National Survey on Drug Use and Health. In S. B. Cohen & J. M. Lepkowski (Eds.), *Eighth Conference on Health Survey Research Methods* (DHHS Publication No. PHS 04-1013, pp. 73-78). Hyattsville, MD: U.S. Department of Health and Human Services, Public Health Service, Centers for Disease Control and Prevention, National Center for Health Statistics.

PURPOSE/OVERVIEW: Response rates traditionally have been highest among the youngest respondents and lowest among the oldest, with the lowest rates found in the 50 or older (50+) age group. The introduction in 2002 of a series of methodological enhancements to the National Survey on Drug Use and Health (NSDUH) appeared to improve the response rates for most age groups but had only a small impact on the 50+ age group (Kennet, Gfroerer, Bowman, Martin, & Cunningham, 2003). Because lower response rates make nonresponse bias more likely, and because there is a disturbingly low response rate among the 50+ group, this paper aims to understand why this may be in order to understand how the problem can be ameliorated. This topic is of increasing importance as the proportion of Americans in this age group increases (U.S. Bureau of the Census, 1999). Obtaining unbiased survey estimates will be vital to assess accurately the substance abuse treatment need for older Americans in the coming years. This need is expected to nearly triple by 2020 as the baby boom generation carries its alcohol and drug use into older ages (Gfroerer, Penne, Pemberton, & Folsom, 2002). The purpose of this paper is to provide a better understanding of nonresponse among older sample members in NSDUH in order to tailor methods to improve response rates and reduce the threat of nonresponse error.

METHODS: This paper examines the components of nonresponse (refusals, noncontacts, and/or other incompletes) among the 50+ age group in NSDUH. It also examines respondent, environmental, and interviewer characteristics in order to identify the correlates of nonresponse among the 50+ group, including relationships that are unique to the 50+ group. Finally, this paper considers the root causes for differential nonresponse by age, drawing from focus group sessions with NSDUH field interviewers on the topic of nonresponse among the 50+ group.

RESULTS/CONCLUSIONS: This paper showed that nonresponse in NSDUH was higher among the 50+ group than among any other age group and was primarily due to a high rate of refusals, especially among sample members aged 50 to 69, and a high rate of physical and mental incapability among those 70 or older. Taken together with evidence from interviewer focus groups, it appeared that the higher rate of refusal among the 50+ age group may, in part, have

been due to fears and misperceptions about the survey and interviewers' intentions. Increased public awareness about the study may allay these fears. Although an increase in the incentive amount may not automatically increase response rates among this group, the authors concluded that other protocol changes and methodological enhancements may be effective.

Imputation and unbiased estimation: Use of the centered predictive mean neighborhoods method

CITATION: Singh, A., Grau, E., & Folsom, R. (2004). Imputation and unbiased estimation: Use of the centered predictive mean neighborhoods method. In *Proceedings of the 2004 Joint Statistical Meetings, American Statistical Association, Section on Survey Research Methods, Toronto, Ontario, Canada* [CD-ROM]. Alexandria, VA: American Statistical Association.

PURPOSE/OVERVIEW: Methods for determining the predictive distribution for multivariate imputation range between two extremes, both of which are commonly employed in practice: a completely parametric model-based approach, and a completely nonparametric approach, such as the nearest neighbor hot deck (NNHD). A semiparametric middle ground between these two extremes is to fit a series of univariate models and construct a neighborhood based on the vector of predictive means. This is what is done under the predictive mean neighborhoods (PMN) method, a generalization of Rubin's predictive mean matching method. Because the distribution of donors in the PMN neighborhood may not be centered at the recipient's predictive mean, estimators of population means and totals could be biased. To overcome this problem, the authors propose a modification to PMN that uses sampling weight calibration techniques, such as the generalized exponential model (GEM) method of Folsom and Singh to center the empirical distribution from the neighborhood.

METHODS: Empirical results on bias and mean squared error (MSE), based on a simulation study using data from the 2002 National Survey on Drug Use and Health, are presented to compare the centered PMN with other methods.

RESULTS/CONCLUSIONS: Although it had been theorized that bias could be a problem for existing methods, this simulation study was unable to detect any meaningful bias with any of the methods. Furthermore, none of the methods showed a consistent pattern of higher or lower variance beyond what was expected.

Combined-year state-level public use files and single-year nation-level PUFs from the National Survey of [sic] Drug Use and Health (NSDUH) data

CITATION: Wright, D., & Singh, A. (2004). Combined-year state-level public use files and single-year nation-level PUFs from the National Survey of [sic] Drug Use and Health (NSDUH) data. In *Proceedings of the 2004 Joint Statistical Meetings, American Statistical Association, Section on Survey Research Methods, Toronto, Ontario, Canada* [CD-ROM]. Alexandria, VA: American Statistical Association.

PURPOSE/OVERVIEW: Since 1999, the Substance Abuse and Mental Health Services Administration (SAMHSA) has provided yearly national public use files (PUFs) for the National Survey on Drug Use and Health (NSDUH) data using a procedure based on the micro agglomeration, substitution, subsampling, and calibration (MASSC) system for statistical disclosure limitation. There is a growing demand for State-level data, and SAMHSA is considering providing State-level PUFs based on combining several years of NSDUH data. The authors explore various concerns and approaches to State-level PUFs and indicated how MASSC could address some of them.

METHODS: Releasing combined-year State-level PUFs alongside single-year national PUFs poses several challenges. The most important one is that confidentiality of an individual could be compromised if an intruder were able to match the State-level PUFs with the national PUFs on the basis of various sensitive variables that are typically not perturbed, and thus may succeed in attaching State identifiers to the national PUFs. This problem can be reduced by taking advantage of the randomness in perturbation and suppression used in MASSC.

RESULTS/CONCLUSIONS: In this paper, the authors suggest a way in which State-level PUFs could be created if one is willing to combine data over several years. The State PUFs would provide a wealth of information for each State for the calculation of point estimates and for analyzing relationships within each State.

2005

A test of the item count methodology for estimating cocaine use prevalence

CITATION: Biemer, P. P., Jordan, B. K., Hubbard, M., & Wright, D. (2005). A test of the item count methodology for estimating cocaine use prevalence. In J. Kennet & J. Gfroerer (Eds.), *Evaluating and improving methods used in the National Survey on Drug Use and Health* (DHHS Publication No. SMA 05-4044, Methodology Series M-5, pp. 149-174). Rockville, MD: Substance Abuse and Mental Health Services Administration, Office of Applied Studies.

PURPOSE/OVERVIEW: The Substance Abuse and Mental Health Services Administration (SAMHSA) has long sought ways to improve the accuracy of the prevalence estimates provided by the National Survey on Drug Use and Health (NSDUH). One method of data collection that shows some promise for improving reporting accuracy is the "item count method." This technique provides respondents with an enhanced perception of anonymity when reporting a sensitive behavior, such as drug use. Past experience with the item count (IC) methodology (e.g., Droitcour et al., 1991) has identified two major problems with this method: task difficulty and the selection of the innocuous items for the list.

METHODS: To test the efficacy of the IC methodology for estimating drug use prevalence, the method was implemented in the 2001 survey. This chapter describes the research conducted in 2000 and 2001 to (1) develop an IC module for past year cocaine use, (2) evaluate and refine the module using cognitive laboratory methods, (3) develop the prevalence estimators of past year cocaine use for the survey design, and (4) make final recommendations on the viability of using the IC method to estimate drug use prevalence. As part of item (4), IC estimates of past year cocaine use based on this implementation are presented, and the validity of the estimates is discussed.

RESULTS/CONCLUSIONS: Considerable effort was directed toward the development, implementation, and analysis of an IC methodology for the estimation of cocaine use prevalence in NSDUH. Several adaptations of existing methods were implemented, offering hope that the refined method would succeed in improving the accuracy of prevalence estimation. Despite these efforts, the IC methodology failed to produce estimates of cocaine use that were even at the level of those obtained by simply asking respondents directly about their cocaine use. Because the direct questioning method is believed to produce underestimates of cocaine, these findings suggest that the IC methodology is even more biased than self-reports.

Evaluation of follow-up probes to reduce item nonresponse in NSDUH

CITATION: Caspar, R. A., Penne, M. A., & Dean, E. (2005). Evaluation of follow-up probes to reduce item nonresponse in NSDUH. In J. Kennet & J. Gfroerer (Eds.), *Evaluating and improving methods used in the National Survey on Drug Use and Health* (DHHS Publication No. SMA 05-4044, Methodology Series M-5, pp. 121-148). Rockville, MD: Substance Abuse and Mental Health Services Administration, Office of Applied Studies.

PURPOSE/OVERVIEW: Technological advances in survey research since 1990 offer the ability to improve the quality of face-to-face survey data in part by reducing item nonresponse. Assuming the instrument has been programmed correctly, computer-assisted interviewing (CAI) can eliminate missing data caused by confusing skip instructions, hard-to-locate answer spaces, and simple inattention to the task at hand. Yet in and of itself, CAI cannot reduce the item nonresponse created when a respondent chooses to give a "Don't know" or "Refused" response to a survey question. Previous research (see, e.g., Turner, Lessler, & Gfroerer, 1992) has shown that these types of responses are more common in self-administered questionnaires than in those administered by an interviewer. Most likely, this occurs because in a self-administered interview the interviewer does not see the respondent's answer and thus cannot probe or follow up on these types of incomplete responses. This chapter introduces a methodology designed to reduce item nonresponse to critical items in the audio computer-assisted self-interviewing (ACASI) portion of the questionnaire used in the National Survey on Drug Use and Health (NSDUH). Respondents providing "Don't know" or "Refused" responses to items designated as essential to the study's objectives received tailored follow-up questions designed to simulate interviewer probes.

METHOD: The results are based on an analysis of unweighted data collected for the 2000 survey (n = 71,764). In total, 2,122 respondents (3.0 percent) triggered at least one of the 38-item nonresponse follow-up questions in 2000. Demographic characteristics are provided of those respondents who triggered at least one follow-up item. To determine what other respondent characteristics tend to be associated with triggering follow-up questions, several multivariate models were developed. Logistic regression was used to determine the likelihood of triggering a follow-up (e.g., answering "Don't know" or "Refused" to a critical lifetime, recency, or frequency drug question). The lifetime follow-up model was run on all cases, while the recency and frequency follow-up models were run only on those subsets of respondents who reported lifetime use (in the case of the recency follow-up item) or who reported use in the past 30 days (in the case of the frequency follow-up). The predictor variables included in these models are described.

RESULTS/CONCLUSIONS: Perhaps the most significant finding from these analyses is that item nonresponse to these critical items was quite low. For the most part, respondents were willing to answer these questions and did not require additional prompting to do so. As a result of the low item nonresponse rates, the data presented must be interpreted with care. The results of these analyses suggest that younger respondents were more likely to trigger the follow-ups and to provide substantive responses to the follow-ups. In addition, the follow-up methodology was more successful in converting respondents who triggered the follow-up through a "Don't know" response than through a "Refused" response. The methodology also was more successful when combined with a revised question that reduced respondent recall burden, as was done with the 30-day frequency follow-up for "Don't know." The largest percentage of follow-up responders provided a substantive response to the 30-day frequency follow-up when the question was simplified by providing response categories in place of the open-ended response field. There also was some evidence to suggest that drug use may be more prevalent among the follow-up responders although small sample sizes precluded a thorough examination of this result.

Taken together, these results suggest the follow-up methodology is a useful strategy for reducing item nonresponse, particularly when the nonresponse is due to "Don't know" responses.

Additional thought should be given to whether improvements can be made to the refusal follow-ups to increase the number of respondents who convert to a substantive response. Focus groups could be useful in identifying other reasons (beyond the fear of disclosure and questions about the importance of the data) that could cause respondents to refuse these critical items. The results of such focus groups could be used to develop more appropriately worded follow-ups that might be more persuasive in convincing respondents that they should provide substantive responses.

Association between interviewer experience and substance use prevalence rates in NSDUH

CITATION: Chromy, J. R., Eyerman, J., Odom, D., McNeeley, M. E., & Hughes, A. (2005). Association between interviewer experience and substance use prevalence rates in NSDUH. In J. Kennet & J. Gfroerer (Eds.), *Evaluating and improving methods used in the National Survey on Drug Use and Health* (DHHS Publication No. SMA 05-4044, Methodology Series M-5, pp. 59-88). Rockville, MD: Substance Abuse and Mental Health Services Administration, Office of Applied Studies.

PURPOSE/OVERVIEW: Analysis of survey data from the National Survey on Drug Use and Health (NSDUH) has shown a relationship between interviewer experience, response rates, and the prevalence of self-reported substance use (Eyerman, Odom, Wu, & Butler, 2002; Hughes, Chromy, Giacoletti, & Odom, 2001, 2002). These analyses have shown a significant and positive relationship between the amount of prior experience an interviewer has with collecting NSDUH data and the response rates that the interviewer produces with his or her workload. The analyses also have shown a significant and negative relationship between the amount of prior experience of an interviewer and the prevalence of substance use reported in cases completed by that interviewer. This chapter describes the methodology employed to explain these effects within a unified theoretical framework.

METHODS: The prior analyses mentioned above examined interviewer response rates and prevalence independently. This has made it difficult to determine whether the lower prevalence estimates for experienced interviewers are a result of the change in the sample composition due to higher response rates or whether the lower prevalence estimates are a result of a direct effect of interviewer behavior on respondent self-reporting. This study combines these two explanations to produce a conceptual model that summarizes the expectations for the relationship between interviewer experience and prevalence estimates. The combined explanation from the conceptual model is evaluated in a series of conditional models to examine the indirect effect of response rates and the direct effect of interviewer experience on prevalence estimates.

RESULTS/CONCLUSIONS: The analysis shows that increased interviewer experience simultaneously increases response rates and decreases prevalence estimates. In addition, the effect of increased interviewer experience on prevalence cannot be fully explained by weight adjustments based on earlier models (i.e., screening and interview level). In other words, the interviewer effect on prevalence cannot be fully attributed to the increase in response rates by experienced interviewers. Furthermore, interviewer experience was significant in the final model, showing that the covariates also did not account for all the decrease in prevalence. A statistical analysis of marginal and incremental prevalence estimates based on three levels of

interviewer experience showed that plausible explanations for the decrease in prevalence for experienced interviewers include (1) lower substance use reporting by the additional respondents and (2) lower reporting of substance use by respondents that interviewers with all levels of experience interview.

Comparing NSDUH income data with income data in other datasets

CITATION: Cowell, A. J., & Mamo, D. (2005). Comparing NSDUH income data with income data in other datasets. In J. Kennet & J. Gfroerer (Eds.), *Evaluating and improving methods used in the National Survey on Drug Use and Health* (DHHS Publication No. SMA 05-4044, Methodology Series M-5, pp. 175-188). Rockville, MD: Substance Abuse and Mental Health Services Administration, Office of Applied Studies.

PURPOSE/OVERVIEW: Personal income and family or household income are among the many demographic measures obtained in the National Survey on Drug Use and Health (NSDUH). Because income is a correlate of substance use and other behaviors, it is important to evaluate the accuracy of the income measure in NSDUH. One metric of accuracy is to compare the estimated distribution of income based on NSDUH with the distributions from other data sources that are used frequently. This chapter compares the distribution of 1999 personal income data from the 2000 NSDUH with the distributions in the same year from the Current Population Survey (CPS) and the statistics of income (SOI) data.

METHODS: The income measure used from the two surveys (NSDUH and the CPS) is personal income, rather than family or household income. Although the CPS explicitly gathers information about more sources of personal income than NSDUH, the income sources in the CPS were combined to map exactly to NSDUH. The income measure from the SOI is adjusted gross income (AGI), reported by various combinations of filing status. Two sets of comparisons were made that differed by marital status in the surveys and filing status in the SOI. First, the income distribution reported in the surveys, regardless of marital status, was compared with that in the SOI, regardless of filing status. For those tax-filing units whose status was "married filing jointly," the reported AGI was for two people, whereas the survey data used were for individual income only. Consequently, these comparisons were unlikely to provide a close fit of the income distributions, particularly in higher income intervals. Second, the income distribution in the surveys for those who were unmarried (excluding widow[er]s) was compared with the income distribution in the SOI for those whose filing status was single. Because only unmarried people can file as "single," the restrictions in this second set of comparisons should have ensured a relatively close fit between the survey data and the SOI. The data did not allow a reasonable comparison to be made between the income distribution in NSDUH and the distribution for other filing statuses in the SOI, such as "married, filing jointly." Because pair-level weights were not available for NSDUH at the time of this analysis, a reporting unit in NSDUH could not be created so that it could compare reasonably with "married filing jointly" in the SOI. NSDUH weights are calibrated to represent individuals. For NSDUH data to represent a pair of married people, rather than individuals, a different set of weights—pair-level weights—are required.

RESULTS/CONCLUSIONS: Despite some fundamental differences between the SOI and either of the survey datasets (CPS and NSDUH), there were strong similarities between the three

income distributions. In both sets of comparisons, the frequencies reported in NSDUH and the CPS were typically within 2 percentage points of each other across all income intervals. With the exception of the lowest interval, in the second set of comparisons (single people and single filers), the frequencies of the three datasets were within 2.5 percentage points of one another across all income intervals.

The differential impact of incentives on cooperation and data collection costs: Results from the 2001 National Household Survey on Drug Abuse incentive experiment

CITATION: Eyerman, J., Bowman, K., Butler, D., & Wright, D. (2005). The differential impact of incentives on cooperation and data collection costs: Results from the 2001 National Household Survey on Drug Abuse incentive experiment. *Journal for Social and Economic Measurement, 30*(2-3), 157-169.

PURPOSE/OVERVIEW: Research has shown that cash incentives can increase cooperation rates and response rates in surveys. The purpose of this paper is to determine whether the impact of incentives in cooperation is consistent across subgroups.

METHODS: An experiment on differing levels of incentives was conducted during the 2001 National Household Survey on Drug Abuse (NHSDA). Respondents were randomly assigned to receive either a $40 incentive, a $20 incentive, or no incentive. The results of the experiment were assessed by examining descriptive tables and analyzing logistic regression models.

RESULTS/CONCLUSIONS: Overall, respondents in the incentive group had higher cooperation rates while lowered data collection costs. The increased response rate did not significantly change the population estimates for drug abuse. The results of the logit model revealed different levels of cooperation for different demographic subgroups. However, the incentives neither enhanced nor reduced the difference in levels of cooperation across subgroups. The results indicate that it was beneficial for the survey to use incentives to encourage cooperation.

Development of a Spanish questionnaire for NSDUH

CITATION: Hinsdale, M., Díaz, A., Salinas, C., Snodgrass, J., & Kennet, J. (2005). Development of a Spanish questionnaire for NSDUH. In J. Kennet & J. Gfroerer (Eds.), *Evaluating and improving methods used in the National Survey on Drug Use and Health* (DHHS Publication No. SMA 05-4044, Methodology Series M-5, pp. 89-104). Rockville, MD: Substance Abuse and Mental Health Services Administration, Office of Applied Studies.

PURPOSE/OVERVIEW: Translation of survey questionnaires is becoming standard practice for large-scale data collection efforts as growing numbers of immigrants arrive in the United States. However, the methods used to produce survey translations have not been standardized—even for Spanish, the most common target language (Shin & Bruno, 2003). The translation review of the National Survey on Drug Use and Health (NSDUH) was carried out for a variety of reasons. For many years, the survey has provided a Spanish-language version of the instrument for

respondents who requested it. Each year, as new questions were added to the survey, translations were carried out on an ad hoc basis using a variety of translators. In the 1999 survey redesign, a large number of questions were added, and a large number of existing questions were altered to accommodate the audio computer-assisted self-interviewing (ACASI) format. It became apparent through feedback from the field that some of the Spanish questions seemed awkward; consequently, survey staff decided that a comprehensive review would be appropriate. It was determined that a multicultural review of the 2000 survey's Spanish instrument would be the most effective procedure.

METHODS: This chapter describes the techniques and principles that were applied in a multicultural review of the translation of the NSDUH questionnaire. Common problems that arose in the translation process and "best practices" for their resolution also are discussed. In addition, because increasing numbers of surveys are employing computer-assisted interviewing (CAI), this chapter illustrates some of the ways in which this technology ideally can be put to use in conveying translated materials. Using three translators coming from varied backgrounds in Central and South America, and focus groups of potential respondents representing the major Spanish-speaking groups in the United States, a translation service that specialized in this type of work carried out a review of the entire questionnaire. The specifics of the focus group and multicultural review processes that took place are described in this chapter within the context of a discussion of best practices for the development of Spanish survey translations.

RESULTS/CONCLUSIONS: Several critical steps in the development of accurate Spanish survey translations were identified by the authors. A seemingly obvious first step involves staffing the project with qualified personnel—from translators to interviewers. To optimize respondent comprehension, translations should be developed using a multicultural approach and should be tested and reviewed thoroughly by a diverse group of bilingual persons (Schoua-Glusberg, 2000). Understanding and applying the concept of cognitive equivalence versus direct translation are key in the development of an effective survey translation. Just as in questionnaire development of English-language surveys, cognitive testing should be employed to identify and correct potential flaws in wording. For studies such as NSDUH that use ACASI, a professional Spanish-speaking voice and skilled audio technicians are needed to ensure the high quality of the audio recording, which maximizes respondents' comprehension. Bilingual interviewers should be fluent and literate in both Spanish and English, and these skills must be demonstrated using a standardized certification procedure. Finally, allowing sufficient time to implement the Spanish translation and train the interviewing staff is perhaps the most problematic step of all because data collection schedules are typically rigorous and researchers are often challenged to maintain the critical timeline even without translations.

Forward telescoping bias in reported age of onset: An example from cigarette smoking

CITATION: Johnson, E. O., & Schultz, L. (2005). Forward telescoping bias in reported age of onset: An example from cigarette smoking. *International Journal of Methods in Psychiatric Research, 14*(3), 119-129.

PURPOSE/OVERVIEW: Age at the onset of a disorder is a critical characteristic that may predict the increased risk of a severe course and genetic liability. However, retrospectively reported onset in surveys is subject to measurement error. This article investigates forward telescoping, a bias in which respondents report events closer to the time of interview than is true. Past research suggests that forward telescoping influences reported age at onset of first substance use, but it does not answer other questions, such as "Is there a difference in the influence of forward telescoping on age of initiation between experimental users (those who have ever used drugs or smoke on a regular basis) and those who use drugs regularly?" or "Does forward telescoping affect reported age at onset for more advanced stages of substance use?" Thus, the purpose of this paper is to examine the effect of this bias on age of onset for smoking initiation and daily smoking.

METHODS: To estimate the effect of age at interview independent of birth year cohort based on multiple cross-sectional surveys of the same population, the authors selected respondents born between 1966 and 1977 (n = 82,122) from the 1997–2000 National Household Surveys on Drug Abuse (NHSDA). Logistic regression was used to estimate the magnitude of forward telescoping in reported age when the first cigarette was smoked to test whether forward telescoping was greater for experimental smokers than for regular smokers and to assess whether the magnitude of forward telescoping in reported age of first daily smoking was lower than that of initiation.

RESULTS/CONCLUSIONS: The authors found an association between age at onset and age at interview, within birth year, for experimenters and for daily smokers. In addition, as age at interview increased from 12 to 25, the authors found that the probability of reporting early onset decreased by half. Contrary to the hypothesis, forward telescoping of age at initiation appeared to affect equally both experimental smokers and daily smokers. However, it also was found that the degree of forward telescoping of age at initiation of smoking differed significantly by gender and that significant forward telescoping of age at onset of daily smoking occurred differentially by race/ethnicity. Overall, the results of the analysis suggest that forward telescoping is a nonignorable bias that can possibly be mitigated through attention to components of the survey design process, such as question and sample design.

Evaluating and improving methods used in the National Survey on Drug Use and Health

CITATION: Kennet, J., & Gfroerer, J. (Eds.). (2005). *Evaluating and improving methods used in the National Survey on Drug Use and Health* (DHHS Publication No. SMA 05-4044, Methodology Series M-5). Rockville, MD: Substance Abuse and Mental Health Services Administration, Office of Applied Studies.

PURPOSE/OVERVIEW: The National Survey on Drug Use and Health (NSDUH) is the leading source of information on the prevalence and incidence of the use of alcohol, tobacco, and illicit drugs in the United States. It is currently administered to approximately 67,500 individuals annually, selected from the civilian, noninstitutionalized population of the United States, including Alaska and Hawaii. A survey of this size and importance is compelled to utilize the latest and best methodology over all facets of its operations. Given the sample size and the careful sampling procedures employed, NSDUH provides fertile soil for the testing and

evaluation of new methodologies. Evaluation of NSDUH methodologies has been and continues to be an integral component of the project. This includes not only reviewing survey research literature and consulting with leading experts in the field, but also conducting specific methodological studies tailored to the particular issues and problems faced by this survey (Gfroerer, Eyerman, & Chromy, 2002; Turner, Lessler, & Gfroerer, 1992).

METHODS: This volume provides an assortment of chapters covering some of the recent methodological research and development in NSDUH, changes in data collection methods and instrument design, as well as advances in analytic techniques. As such, it is intended for readers interested in particular aspects of survey methodology and is a must-read for those with interests in analyzing data collected in recent years by NSDUH.

RESULTS/CONCLUSIONS: This volume contains a collection of some recent methodological research carried out under the auspices of the NSDUH project. Publishing these studies periodically provides a resource for survey researchers wishing to catch up on the latest developments from this unique survey. Readers from a variety of backgrounds and perspectives will find these chapters interesting and informative and, it is hoped, useful in their own careers in survey methodological research, drug abuse prevention and treatment, and other areas.

Introduction. In J. Kennet & J. Gfroerer (Eds.), Evaluating and improving methods used in the National Survey on Drug Use and Health

CITATION: Kennet, J., & Gfroerer, J. (2005). Introduction. In J. Kennet & J. Gfroerer (Eds.), *Evaluating and improving methods used in the National Survey on Drug Use and Health* (DHHS Publication No. SMA 05-4044, Methodology Series M-5, pp. 1-6). Rockville, MD: Substance Abuse and Mental Health Services Administration, Office of Applied Studies.

PURPOSE/OVERVIEW: The National Survey on Drug Use and Health (NSDUH) is the leading source of information on the prevalence and incidence of the use of alcohol, tobacco, and illicit drugs in the United States. It is currently administered to approximately 67,500 individuals annually, selected from the civilian, noninstitutionalized population of the United States, including Alaska and Hawaii. A survey of this size and importance is compelled to utilize the latest and best methodology over all facets of its operations. Given the sample size and the careful sampling procedures employed, NSDUH provides fertile soil for the testing and evaluation of new methodologies. Evaluation of NSDUH methodologies has been and continues to be an integral component of the project. This includes not only reviewing survey research literature and consulting with leading experts in the field, but also conducting specific methodological studies tailored to the particular issues and problems faced by this survey (Gfroerer, Eyerman, & Chromy, 2002; Turner, Lessler, & Gfroerer, 1992). This volume provides an assortment of chapters covering some of the recent methodological research and development in NSDUH, changes in data collection methods and instrument design, as well as advances in analytic techniques.

METHODS: This introduction begins with a brief history and description of NSDUH. A more detailed account can be found in Gfroerer et al. (2002). Prior methodological research on

NSDUH then is described, followed by an account of the major methodological developments that were implemented in 2002. Finally, each of the chapters and their authors are introduced.

RESULTS/CONCLUSIONS: This volume contains a collection of some recent methodological research carried out under the auspices of the NSDUH project. Publishing these studies periodically provides a resource for survey researchers wishing to catch up on the latest developments from this unique survey. Readers from a variety of backgrounds and perspectives will find these chapters interesting and informative and, it is hoped, useful in their own careers in survey methodological research, drug abuse prevention and treatment, and other areas.

Introduction of an incentive and its effects on response rates and costs in NSDUH

CITATION: Kennet, J., Gfroerer, J., Bowman, K. R., Martin, P. C., & Cunningham, D. (2005). Introduction of an incentive and its effects on response rates and costs in NSDUH. In J. Kennet & J. Gfroerer (Eds.), *Evaluating and improving methods used in the National Survey on Drug Use and Health* (DHHS Publication No. SMA 05-4044, Methodology Series M-5, pp. 7-18). Rockville, MD: Substance Abuse and Mental Health Services Administration, Office of Applied Studies.

PURPOSE/OVERVIEW: In 2002, the National Survey on Drug Use and Health (NSDUH) began offering respondents a $30 cash incentive for completing the questionnaire. This development occurred within the context of a name change and other methodological improvements to the survey and resulted in significantly higher response rates. Moreover, the increased response rates were achieved in conjunction with a net decrease in costs incurred per completed interview. This chapter presents an analysis of response rate patterns by geographic and demographic characteristics, as well as interviewer characteristics, before and after the introduction of the incentive. Potential implications for other large-scale surveys also are discussed.

METHODS: To demonstrate the effects of the incentive and other changes, a comparison is presented of the response rates, by quarters, in the years before and after the incentive was introduced. These rates then are broken down by geographic area, population density, respondent age, gender, and race/ethnicity. Response rates also are examined with respect to the characteristics of interviewers, including their prior experience on the survey, race/ethnicity, and gender.

RESULTS/CONCLUSIONS: The $30 incentive, with possible help from the other changes that were introduced in January 2002, produced dramatic improvement in the number of eligible respondents who agreed to complete the NSDUH interview. Moreover, the increase in respondent cooperation was accompanied by a decrease in cost per completed interview. Clearly, the adoption of the incentive was beneficial to all involved, with the possible exception of the field interviewers, who required fewer hours to complete their assignments on the project and consequently received less pay.

From these analyses, it appears that incentives had their most pronounced effect among people between the ages of 12 and 25. Because these are known to be the years in which substance use and abuse are most prevalent and have their greatest incidence, it seems the incentive was well spent in terms of capturing the population of most interest. However, serious thought needs to be given to methods for attracting those older than 25. It could be the case that $30 was simply insufficient to attract people who have settled into careers and/or other more rewarding activities, such as child rearing or retirement, or it could be that these people did not participate for other reasons. These reasons will have to be investigated and addressed in order for NSDUH to optimally cover the aging baby-boom generation and other cohorts.

The use of monetary incentives in federal surveys on substance use and abuse

CITATION: Kulka, R. A., Eyerman, J., & McNeeley, M. E. (2005). The use of monetary incentives in federal surveys on substance use and abuse. *Journal for Social and Economic Measurement, 30*(2-3), 233-249.

PURPOSE/OVERVIEW: Empirical research over the past 30 years has shown positive effects of monetary incentives on response rates. This paper discusses the use of incentives specifically for surveys on substance use and abuse.

METHODS: This paper starts by providing a background and review of the current empirical literature on the effects that monetary incentives have on survey response rates, survey statistics, and other practical and operational issues in surveys. Next, two controlled experiments on the effect of incentives on substance use surveys are discussed: the Alcohol and Drug Services Study (ADSS) and the National Household Survey on Drug Abuse (NHSDA). Each of the studies randomized respondents into different incentive categories, including no incentives and two to three levels of increasing incentives.

RESULTS/CONCLUSIONS: The ADSS results revealed that higher incentive payments correlated with higher cooperation rates, but that the difference in cooperation rates between the levels of incentive payments was small. The analyses also revealed that the incentive payments did not affect different subgroups differently. In addition, the results indicated that the use of incentives did not affect the quality of respondents' answers. The NHSDA results, however, showed that although incentives did increase response rates, they had differing impacts on different subgroups. The results of the analyses on survey response bias were inconclusive. The results of these two studies revealed that more research is needed on this topic to further understand the effect of incentives on survey data quality. The results of these studies are described in more detail in a series of papers appearing in the *Journal of Social and Economic Measurement*.

Effects of the September 11, 2001, terrorist attacks on NSDUH response rates

CITATION: McNeeley, M. E., Odom, D., Stivers, J., Frechtel, P., Langer, M., Brantley, J., Painter, D., & Gfroerer, J. (2005). Effects of the September 11, 2001, terrorist attacks on NSDUH response rates. In J. Kennet & J. Gfroerer (Eds.), *Evaluating and improving methods*

used in the National Survey on Drug Use and Health (DHHS Publication No. SMA 05-4044, Methodology Series M-5, pp. 31-58). Rockville, MD: Substance Abuse and Mental Health Services Administration, Office of Applied Studies.

PURPOSE/OVERVIEW: The purpose of this study was to determine whether the events of September 11th had an effect on screening response rates (SRRs) and interview response rates (IRRs).

METHODS: In addition to the preliminary statistical analysis of response rate data, logistic regression models were run to provide a more refined analysis. Field interviewer (FI) focus groups also were used to assess changes in the logistics of FI activity and in the use of the lead letter. To capture heightened concerns about security, FIs also were asked about increases in controlled access problems and changes in the mode of contact with screening and interview respondents.

RESULTS/CONCLUSIONS: It was found that the New York City consolidated metropolitan statistical area (CMSA) and Washington, DC, primary metropolitan statistical area (PMSA) response rates suffered dramatic decreases following the September 11th terrorist attacks, although the differences in IRR in the Washington, DC, PMSA were shown to be significant only after modeling on a number of factors. The national screening response rate (SRR) also showed a decrease even after removing the New York City CMSA and Washington, DC, PMSA from the sample. This decrease was significant but less dramatic than in the two metropolitan areas.

Appendix C: Research on the impact of changes in NSDUH methods

CITATION: Office of Applied Studies. (2005). Appendix C: Research on the impact of changes in NSDUH methods. In *Results from the 2004 National Survey on Drug Use and Health: National findings* (DHHS Publication No. SMA 05-4062, NSDUH Series H-28, pp. 145-154). Rockville, MD: Substance Abuse and Mental Health Services Administration.

PURPOSE/OVERVIEW: Although the design of the 2002 through 2004 National Surveys on Drug Use and Health (NSDUH) was similar to the design of the 1999 through 2002 surveys, there were important methodological differences between the 2002 to 2004 NSDUHs and prior surveys, including a change in the survey's name, the introduction of an incentive, improved data collection quality control procedures and the use of the 2000 decennial census for sample weighting. The results of the 2002 survey suggested that the incentive had an impact on estimates. A panel of survey methodology experts concluded that it would not be possible to measure the effects of each changes separately because of the multiple changes made to the survey simultaneously and recommended that the Substance Abuse and Mental Health Services Administration (SAMHSA) continue its analyses of the 2001 and 2002 data to learn as much as possible about the impacts of each of the methodological improvements. The purpose of this appendix is to summarize the studies of the effects of the 2002 method changes and to discuss the implications of this body of research for analysis of NSDUH data.

METHODS: Early analyses were presented to a panel of survey design and survey methodology experts convened on September 12, 2002. The analyses included (1) retrospective cohort analyses; (2) response rate pattern analyses; (3) response rate impact analyses; (4) analyses of the impact of new census data; and (5) model-based analyses of protocol changes, name change, and incentives. Since 2002, two additional analyses were conducted that extend those described above: more in-depth incentive experiment analyses and further analysis of the 2001 field interventions.

RESULTS/CONCLUSIONS: A summary of all of the results of the 2002 NSDUH method analyses was presented to a second panel of consultants on April 28, 2005. The panel concluded that there was no possibility of developing a valid direct adjustment method for the NSDUH data and that SAMHSA should not compare 2002 and later estimates with 2001 and earlier estimates for trend assessment. The panel suggested that SAMHSA make this recommendation to users of NSDUH data.

Analysis of NSDUH record of call data to study the effects of a respondent incentive payment

CITATION: Painter, D., Wright, D., Chromy, J. R., Meyer, M., Granger, R. A., & Clarke, A. (2005). Analysis of NSDUH record of call data to study the effects of a respondent incentive payment. In J. Kennet & J. Gfroerer (Eds.), *Evaluating and improving methods used in the National Survey on Drug Use and Health* (DHHS Publication No. SMA 05-4044, Methodology Series M-5, pp. 19-30). Rockville, MD: Substance Abuse and Mental Health Services Administration, Office of Applied Studies.

PURPOSE/OVERVIEW: The National Survey on Drug Use and Health (NSDUH) employs a multistage area probability sample to produce population estimates of the prevalence of substance use and other health-related issues. Letters are sent to selected households to alert potential respondents of the interviewer's future visit. Interviewers then visit the residence to conduct a brief screening, which determines whether none, one, or two persons are selected from the household. To maximize response rates, interviewers may make several visits to a household to obtain cooperation (the term "call" is used for "visits" from this point on with the understanding that this is a face-to-face survey; telephones are not used to contact potential respondents). In-person interviews are conducted with selected respondents using both computer-assisted personal interviewing (CAPI) and audio computer-assisted self-interviewing (ACASI). Sensitive questions, such as those on illicit drug use, are asked using the ACASI method to encourage honest reporting. A detailed description of the NSDUH methodology is described elsewhere (RTI International, 2003). During the late 1990s, NSDUH experienced a slight decline in response rates. A closer examination of the data revealed stable noncontact patterns, but increasing refusal rates (Office of Applied Studies [OAS], 2001). This implied that sample members were beginning to become less likely to participate once they were contacted. This was compounded by the need to hire a large number of new interviewers who may not have had the confidence or skills to overcome respondent refusals.

METHODS: Given the slight decline in response rates, and the expectation that this trend might become more serious, NSDUH staff designed an experiment to evaluate the effectiveness of

monetary incentives in improving respondent cooperation. A randomized, split-sample, experiment was conducted during the first 6 months of data collection in 2001. The experiment was designed to compare the impact of $20 and $40 incentive treatments with a $0 control group on measures of respondent cooperation and survey costs.

RESULTS/CONCLUSIONS: The results showed that both the $20 and $40 incentives increased overall response rates while producing significant cost savings when compared with the $0 control group (Eyerman, Bowman, Butler, & Wright, 2002). Preliminary analysis showed no statistically detectable effects of the three incentive treatments on selected substance use estimates. Subsequent analysis showed some positive and negative effects depending on the substance use measure when the $20 and $40 treatments were combined and compared with the $0 control group (Wright, Bowman, Butler, & Eyerman, 2002). Based on the outcome of the 2001 experiment, NSDUH staff implemented a $30 incentive payment in 2002. Their analysis showed that a $30 incentive would strike a balance between gains in response rates and cost savings. This chapter analyzes the effect of the new $30 incentive on the data collection process as measured by record of calls (ROC) information. The effect of the incentives implemented in 2002 on response rates and costs is discussed by Kennet et al. in Chapter 2 of this volume.

Analyzing audit trails in NSDUH

CITATION: Penne, M. A., Snodgrass, J., & Barker, P. (2005). Analyzing audit trails in NSDUH. In J. Kennet & J. Gfroerer (Eds.), *Evaluating and improving methods used in the National Survey on Drug Use and Health* (DHHS Publication No. SMA 05-4044, Methodology Series M-5, pp. 105-120). Rockville, MD: Substance Abuse and Mental Health Services Administration, Office of Applied Studies.

PURPOSE/OVERVIEW: For the National Survey on Drug Use and Health (NSDUH), the interview sections on substance use and other sensitive topics were changed in 1999 from self-administered paper-and-pencil interviewing (PAPI) to audio computer-assisted self-interviewing (ACASI). These changes were prompted by research that showed that computer-assisted interviewing (CAI) questionnaires reduced input errors. Research also showed that use of ACASI increased comprehension for less literate respondents and, by increasing privacy, resulted in more honest reporting of illicit drug use and other sensitive behaviors.

METHOD: In this chapter, the earlier work is briefly described, possible methods for streamlining the data-processing portion are discussed, and the use of audit trails in the 2002 survey is focused on to investigate three aspects of data quality: question timing, respondent breakoffs, and respondent "backing up" to change prior responses.

RESULTS/CONCLUSIONS: Timing data showed that when measured against a gold standard (GS) time, field interviewers (FIs) were spending approximately the correct amount of time with the very beginning of the interview at the introduction to the CAI instrument screen. However, once past this point, they spent less time than the GS on several important aspects of the questionnaire, such as setting up the calendar, setting up the ACASI tutorial, completing the verification form, and ending the interview with the respondent. Conversely, they were taking longer than the GS in ending the ACASI portion of the interview.

Decomposing the total variation in a nested random effects model of neighborhood, household, and individual components when the dependent variable is dichotomous: Implications for adolescent marijuana use

CITATION: Wright, D., Bobashev, G. V., & Novak, S. P. (2005). Decomposing the total variation in a nested random effects model of neighborhood, household, and individual components when the dependent variable is dichotomous: Implications for adolescent marijuana use. *Drug and Alcohol Dependence, 78*(2), 195-204.

PURPOSE/OVERVIEW: This paper explores a three-level model used to measure the effect of individual-, household-, and neighborhood-level impacts on reported marijuana use.

METHODS: The model was applied to the 1999 National Household Survey on Drug Abuse (NHSDA) data of approximately 67,000 respondents. The primary outcome variables were a direct and an indirect measure of marijuana use. The multilevel model allowed for individual components that affect drug use and preferences to be more easily identified. However, using this model, it was challenging to discern which component held the most importance. The paper analyzes a method presented by Golstein and Rasbash (1996) for allowing the results to be interpreted more logically and meaningfully.

RESULTS/CONCLUSIONS: Prevalence estimates for past year marijuana use for adolescents, risk of marijuana use, and ease of obtaining drugs are presented. For all three outcomes, the individual-level predictor contributed the most, explaining approximately 80 percent of the variation. The household-level predictors contributed about 15 percent, and the neighborhood-level indicators only about 5 percent. This methodology allowed for the variance estimation of several factors on both binary and continuous measures of drug use.

Non-response bias from the National Household Survey on Drug Abuse incentive experiment

CITATION: Wright, D., Bowman, K., Butler, D., & Eyerman, J. (2005). Non-response bias from the National Household Survey on Drug Abuse incentive experiment. *Journal for Social and Economic Measurement, 30*(2-3), 219-231.

PURPOSE/OVERVIEW: In a preliminary experiment conducted in the 2001 National Household Survey on Drug Abuse (NHSDA), it was concluded that providing incentives increased response rates; therefore, a $30 incentive was used in the subsequent 2002 NHSDA. The purpose of this paper is to explore the effect that incentive had on nonresponse bias.

METHODS: The sample data were weighted by likelihood of response between the incentive and nonincentive cases. Next, a logistic regression model was fit using substance use variables and controlling for other demographic variables associated with either response propensity or drug use.

RESULTS/CONCLUSIONS: The results indicate that for past year marijuana use, the incentive is either encouraging users to respond who otherwise would not respond, or it is encouraging

respondents who would have participated without the incentive to report more honestly about drug use. Therefore, it is difficult to determine whether the incentive money is reducing nonresponse bias, response bias, or both. However, reports of past year and lifetime cocaine did not increase in the incentive category, and past month use of cocaine actually was lower in the incentive group than in the control group.

Author and Subject Index

Author Index

Subject Index