

Alcohol Epidemiologic Data System
Division of Epidemiology and Prevention Research
National Institute on Alcohol Abuse and Alcoholism

Alcohol Epidemiologic Data Directory

September
2022

CSR, Incorporated
22375 Broderick Drive, Suite 220
Sterling, Virginia 20166

Developed Under Contract No. HHSN275201800004C

TABLE OF CONTENTS

| | |
|--|----|
| INTRODUCTION | 4 |
| Section 1: National Health and Alcohol Data Sets..... | 5 |
| Behavioral Risk Factor Surveillance System (BRFSS)—1984–2020, Annually | 7 |
| Collaborative Initiative on Fetal Alcohol Spectrum Disorders (CIFASD)—2003–2007, 2007– 2012, 2012–2017 | 9 |
| Collaborative Study on the Genetics of Alcoholism (COGA)—1991–2016, Ongoing..... | 11 |
| Coronary Artery Risk Development in Young Adults (CARDIA)—1985–1986, 1987–1988, 1990–1991, 1992–1993, 1995–1996, 2000–2001, 2005–2006, 2010–2011, and 2015– 2016 | 13 |
| Drug Abuse Warning Network (DAWN)—1993–2003, 2004–2011, 2018–2021, Annually..... | 15 |
| Fatality Analysis Reporting System (FARS)—1975–2020, Annually | 17 |
| Healthcare Cost and Utilization Project (HCUP) Kids’ Inpatient Database (KID)—1997– 2012, Triennially, and 2016 and 2021 | 18 |
| Healthcare Cost and Utilization Project (HCUP) Nationwide Ambulatory Surgery Sample (NASS)—2016–2019, Annually..... | 19 |
| Healthcare Cost and Utilization Project (HCUP) Nationwide Emergency Department Sample (NEDS)—2006–2019, Annually | 21 |
| Healthcare Cost and Utilization Project (HCUP) National (Nationwide) Inpatient Sample (NIS)—1988–2019, Annually | 22 |
| Healthcare Cost and Utilization Project (HCUP) National Readmissions Database (NRD)— 2010–2018, Annually | 24 |
| Health Information National Trends Survey (HINTS)—2003, 2005, 2007, 2009, 2011–2015, and 2017–2020, Annually..... | 25 |
| National Alcohol Survey (NAS)—1964–1965, 1967, 1969, 1974, 1979, 1984, 1990, 1992, 1995– 1996, 2000–2001, 2005, 2009–2010, and 2014–2015..... | 27 |
| National Ambulatory Medical Care Survey (NAMCS)—1973–1992, 1993–2016, 2018, Annually | 28 |
| National Automotive Sampling System General Estimates System (NASS GES)— 1988– 2015, Crash Report Sampling System (CRSS)—2016–2020, Annually | 29 |
| National COVID Cohort Collaborative (N3C) Data Enclave—2020–2022, Ongoing | 30 |
| National Crime Victimization Survey (NCVS)—1973–2020, Annually | 32 |
| National Emergency Medical Services Information System (NEMSIS)—2009–2020, Annually | 33 |
| National Epidemiologic Survey on Alcohol and Related Conditions (NESARC)—Wave 1 (2001–2002), Wave 2 (2004–2005), Wave 3 (2012–2013) | 34 |
| National Health and Nutrition Examination Survey I (NHANES I)—1971–1975 | 36 |
| National Health and Nutrition Examination Survey I Epidemiologic Follow-up Studies (NHEFS82, NHEFS86, NHEFS87, NHEFS92)..... | 38 |
| National Health and Nutrition Examination Survey II (NHANES II)—1976–1980 | 40 |
| National Health and Nutrition Examination Survey III (NHANES III)—1988–1994 | 41 |
| National Health and Nutrition Examination Survey (Continuous NHANES)— 1999–2020, Biennially | 42 |
| National Health Interview Survey (NHIS)—General Description, 1997–2021, Annually | 43 |
| National Hospital Ambulatory Medical Care Survey (NHAMCS)—1992–2019, Annually | 45 |
| National Hospital Care Survey (NHCS)—2013–2016, Annually, and 2020–2021 | 46 |
| National Hospital Discharge Survey (NHDS)—1970–2010, Annually | 48 |
| National Survey on Drug Use and Health (NSDUH)—2002–2020, Annually..... | 50 |
| National Mental Health Services Survey (N-MHSS)—2014–2020, Biennially; National Substance Use and Mental Health Services Survey (N-SUMHSS)—2021..... | 52 |
| National Survey of Children’s Exposure to Violence (NatSCEV)—1990–2008, 2011, and 2014..... | 54 |
| National Survey of Drinking and Driving Attitudes and Behavior—1991, 1993, 1995, 1997, 1999, 2001, 2004, and 2008..... | 56 |
| National Survey of Family Growth (NSFG)—1973, 1976, 1982, 1988, 1995, 2002, 2006–2010, 2011–2013, 2013–2015, 2017–2019 | 57 |

| | |
|---|-----------|
| National Survey of Substance Abuse Treatment Services (N-SSATS)—2000, 2002–2019, Annually | 58 |
| National Violent Death Reporting System (NVDRS)—2003–2019, Annually | 60 |
| Panel Study of Income Dynamics (PSID)—1968–1996, Annually, 1997–2019, Biennially | 62 |
| Population Assessment of Tobacco and Health (PATH)—2013–2021, Annually | 63 |
| Research and Development Survey (RANDS)—RANDS 1 (2015), RANDS 2 (2016), RANDS 3 (2019)..... | 65 |
| Treatment Episode Data Set (TEDS)—1992–2019, Annually..... | 67 |
| Understanding America Study (UAS)—2014–2018, Ongoing..... | 69 |
| Vital Statistics Mortality Data, Mortality Detail and Multiple Cause of Death —1968–2020, Annually | 71 |
| Section 2: Special Population Data Sets..... | 72 |
| Adolescent Behaviors and Experiences Survey (ABES)—2021..... | 74 |
| American College Health Association-National College Health Assessment (ACHA-NCHA)— Fall 2000–Spring 2008; NCHA II—Fall 2008–Spring 2011; NCHA IIB—Fall 2011– Spring 2015; NCHA IIC—Fall 2015–Spring 2019; NCHA III—Fall 2019–Present | 76 |
| Arrestee Drug Abuse Monitoring (ADAM I)—2000–2003; (ADAM II)—2007–2013, Annually | 78 |
| CIRP Freshman Survey (TFS)—1966–2020, Annually..... | 80 |
| College Senior Survey (CSS)—1993–2020, Annually..... | 81 |
| Department of Defense Health Related Behaviors Survey — 2015 and 2018. Department of Defense Health Behavior Survey—2008 and 2011. Worldwide Surveys of Alcohol and Nonmedical Drug Use Among Military Personnel—1980 and 1982. Worldwide Surveys of Substance Abuse and Health Behaviors Among Military Personnel—1982, 1985, 1988, 1992, 1995, 1998, 2002, and 2005. | 82 |
| The Health and Retirement Study: A Longitudinal Study of Health, Retirement, and Aging (HRS)—1992–2020, Biennially..... | 83 |
| Healthy Minds Study (HMS)—2015–2016, 2016–2017, 2017–2018, 2018–2019, 2019–2020, and 2020–2021..... | 85 |
| Hispanic Community Health Study/Study of Latinos (HCHS/SOL)—2008–2011, 2014–2017..... | 87 |
| Midlife in the United States Study 1, 2, 3 (MIDUS 1, 2, 3)—1995/1996, 2004–2005, and 2013–2014 | 89 |
| Monitoring the Future (MTF): A Continuing Study of American Youth—1975–2021, Annually | 91 |
| National Child Abuse and Neglect Data System, Child File (NCANDS)—1995–2020, Annually | 93 |
| National Longitudinal Study on Adolescent to Adult Health (Add Health)—Wave I (1994– 1995), Wave II (1996), Wave III (2001–2002), Wave IV (2007–2008), Wave V (2016– 2018), Parent Study (2015–2017)..... | 94 |
| National Longitudinal Survey of Youth (NLSY79)—1979–1994, 1995–2018, Biennially | 96 |
| National Longitudinal Survey of Youth 1979 (NLSY79) Child and Young Adult (NLSCYA) Child Sample—1986–2018, Biennially; Young Adult Sample—1994–2018, Biennially | 98 |
| National Longitudinal Survey of Youth (NLSY97)—1997–2011, Annually, 2012–2020, Biennially..... | 100 |
| National Survey of Youth in Custody (NSYC)—2008–09; (NYSC-2)—2012; (NYSC-3)—2017– 2018; National Survey of Youth in Custody, Alternate, Supplemental Survey on Drug and Alcohol Use (NYSC-A)—2008–2009, 2012, and 2018 | 101 |
| NEXT Generation Longitudinal Study (NEXT Generation Health Study)—2009/2010– 2016/2017 | 103 |
| Nurses’ Health Study (NHS)—1976–1988, Biennially; and 1989–2019, Annually | 105 |
| The Pregnancy Risk Assessment Monitoring System (PRAMS)—Phase I (1988–1989), Phase II (1990–1995), Phase III (1996–1999), Phase IV (2000–2003), Phase V (2004–2008), Phase VI (2009–2011), Phase VII (2012–2015), and Phase VIII (2016–2020)..... | 107 |
| The Study of Women’s Health Across the Nation (SWAN)—1996–2008..... | 108 |
| Survey of Inmates in State and Federal Correctional Facilities—1974, 1979, 1986, 1991, | |

Alcohol Epidemiologic Data Directory

| | |
|---|-----|
| 1997, and 2004; Survey of Prison Inmates—2016..... | 110 |
| Youth Risk Behavior Survey (YRBS)—1991–2019 (High School), Biennially, 1998 (Alternative High School), 1995 (College), and 1992 (NHIS) | 111 |
| Section 3: AEDS Publications and Products..... | 113 |
| APPENDIX: LIST OF ACRONYMS..... | 117 |

INTRODUCTION

This *Alcohol Epidemiologic Data Directory* (referred to here as the *Directory*) is compiled and updated by the Alcohol Epidemiologic Data System (AEDS), operated by CSR, Incorporated, under contract to the National Institute on Alcohol Abuse and Alcoholism (NIAAA). AEDS's task is to identify, acquire, maintain, and analyze alcohol-related epidemiologic data under the direction of NIAAA's Division of Epidemiology and Prevention Research.

This *Directory* is a current listing of surveys and other relevant data suitable for epidemiologic research on alcohol. Some surveys are designed specifically to answer alcohol-related questions. Other surveys may address different issues but still contain alcohol-related data. The first section includes data sets that are representative of the overall U.S. population, although many use different age categories in the sample design. The second section includes data sets on special populations (e.g., adolescents, prison inmates, military personnel, and older Americans). A final section describes publications and other research products available from AEDS. It is important to note that this *Directory* is not a comprehensive listing of all data sets that are available to researchers. Many small-scale surveys, such as single-state surveys and local area surveys, are excluded, as are data sets that are outdated or not available to the public.

A variety of organizations sponsor or produce these data sets. Each entry includes a source contact as well as internet addresses to assist researchers with obtaining additional information. Data increasingly are available in downloadable formats from the Internet. Information on availability is provided for each data set, including hyperlinks for downloading, when available. The Internet addresses are checked before the *Directory* is published, but as it is likely that some addresses will change over the period of this publication, refer to the source contacts for new Internet addresses if necessary. Unless otherwise specified, these data sets are not available from AEDS but from the listed sponsoring organizations or their contracted providers.

Analytic results from data sets described in this *Directory* often are available on the Internet in tabular or summary form. Further, some data sets can be analyzed online with programs provided by the sponsoring organization. Some useful Internet links include the Inter-university Consortium for Political and Social Research ([ICPSR](#)), Substance Abuse and Mental Health Data Archive ([SAMHDA](#)), the National Archive of Criminal Justice Data ([NACJD](#)), and the National Center for Health Statistics ([NCHS](#)). Links to additional Federal drug data sources are also available through the "Additional Links & Resources" page at <https://www.whitehouse.gov/ondcp/additional-links>. Section 3 describes AEDS publications; they are accessible through NIAAA's website at <https://www.niaaa.nih.gov/>.

An electronic copy of this *Directory* is available at <https://pubs.niaaa.nih.gov/publications/datasys.htm>. AEDS welcomes any suggestions or comments. Direct any comments or requests for additional copies of this or other AEDS publications to:

Alcohol Epidemiologic Data System
CSR, Incorporated
22375 Broderick Drive, Suite 220
Sterling, Virginia 20166
Phone: (703) 312-5220
Fax: (703) 312-5230
Email: AEDSinfo@csrincorporated.com

Section 1:

National Health and Alcohol Data Sets

Behavioral Risk Factor Surveillance System (BRFSS)—1984–2020, Annually

Sponsoring Agency:

BRFSS surveys are conducted by the states and coordinated by the Centers for Disease Control and Prevention (CDC), U.S. Department of Health and Human Services

Contact:

National Center for Chronic Disease Prevention and Health Promotion
CDC

Division of Population Health
4770 Buford Highway, NE, MS S107-6
Atlanta, GA 30341
Telephone: 800-232-4636
<http://www.cdc.gov/brfss/>

Availability:

Data files in SAS transport format are available for download from
https://www.cdc.gov/brfss/annual_data/annual_data.htm.

Overview:

BRFSS is an ongoing data collection program designed to monitor state-level prevalence of the major behavioral risks associated with premature morbidity and mortality among adults. The survey was initiated in 1984, with 15 states participating in the monthly data collection. By 1994, all states and the District of Columbia were participating. Guam, the Virgin Islands, and Puerto Rico were included in 2001–2002. Factors assessed include alcohol and tobacco use, health care coverage, tested for HIV/AIDS, physical activity, and fruit and vegetable consumption. CDC developed standard core questions for states to use to collect data that could be compared across states. The survey also includes many optional modules and state-added questions.

Survey Design/Methodology:

BRFSS is conducted in each participating state on a probability sample of the adult population ages 18 and older. Telephone interviews are conducted during a two-week period each month throughout the year. Most states use a disproportionate stratified sample design. A few states used a Mitofsky-Waksberg design or a simple random sample design. Deviations from sampling frame and weighting protocols exist among states. Initially conducted with paper-administered survey forms, interviews are now conducted through computer-assisted telephone interviewing. Starting in 2011, cell phone surveys were included in the public release data set.

Sample Characteristics:

BRFSS samples vary in size from state to state and from year to year, depending on the number

of states participating and funding availability. In 2020, 401,958 noninstitutionalized adult respondents ages 18 and older from 52 states and territories participated. The BRFSS is designed to collect state-level data, but some regional prevalence estimates are possible from a number of states that stratify their samples.

Alcohol Variables:

Alcohol variables were asked in reference to the past month or the past 30 days, including frequency of consumption and average number of drinks consumed per occasion. Before 2006, binge drinking for both men and women was defined as having 5 or more drinks per occasion; in 2006, the definition for women changed to 4 or more drinks. Variables addressing communication with health professionals regarding alcohol use were included in the 1996–1999 surveys and in select state surveys in 2011. Additional variables included in select survey years include drinking and driving, living with people with alcohol problems, and reducing alcohol consumption for health reasons. Alcohol questions were included in the core questionnaire before 1994. Beginning in 1994, the alcohol section rotated between the core questionnaire and optional modules (Alcohol Consumption, Alcohol Screening & Brief Intervention, Binge Drinking). Eleven states responded to alcohol questions in 1994, all states responded in 1995, 17 in 1996, all in 1997, 12 in 1998, all in 1999, 11 in 2000, 13 in 2003, 10 in 2008, 19 in 2014, 14 in 2017, and 13 in 2019. Five states added their own alcohol questions in 2000. With the exception of Hawaii in 2004 and New Jersey in 2019, all states, the District of

Columbia, and participating U.S. territories responded to the core survey in 2001–2020.

Other Variables:

BRFSS covers demographics, health status, health care access, family planning, asthma, diabetes, oral health, diet, immunization, seatbelt use, history of hypertension, frequency of physical

exercise, amount of recreational activity, access and storage of firearms, mammography, exposure to stress, smoking, women’s health, HIV/AIDS and prevention behaviors (e.g., annual checkups, cancer screening). Substance use variables other than alcohol include marijuana, tobacco, and other drugs. Optional modules allow states to address emerging health issues.

Collaborative Initiative on Fetal Alcohol Spectrum Disorders (CIFASD)—2003–2007, 2007–2012, 2012–2017

Sponsoring Agency:

National Institute on Alcohol Abuse and Alcoholism

Contact:

CIFASD, Consortium Coordinator
6330 Alvarado Court, Suite 100
San Diego, CA 92120
Telephone: 619-594-4566 or Fax: 619-594-1895

Edward Riley, Ph.D.
Telephone: 619-594-4601
eriley@mail.sdsu.edu
<https://cifasd.org/>

Availability:

A data access request form may be downloaded and submitted at <https://cifasd.org/data-sharing>. Applications are reviewed by the CIFASD Data Access Committee.

Overview:

CIFASD is a multidisciplinary, international consortium of research projects and resource cores, established in 2003, and charged with improving prevention, diagnosis, and treatment of FASD. CIFASD addresses issues related to prenatal alcohol exposure across the lifespan using a range of interrelated clinical and preclinical research approaches. CIFASD's goals are to improve identification of people who have been affected by prenatal alcohol exposure by assessing biomarkers and by using neuroimaging techniques; develop novel tools, such as 3D facial and neurobehavioral screening for use in telemedicine; identify risk and resiliency factors; better understand the effects of prenatal alcohol exposure on consequences across the lifespan; and modify an intervention for improving outcomes to reach larger populations of people with FASD and their families. Consortium members are Children's Hospital Los Angeles, Emory University, Helsinki Folkhälsan Research Center, Harvard University, Indiana University Bloomington, Indiana School of Medicine, Indiana-IUPUI School of Science, Moscow Institute of Obstetrics and Gynecology, Rutgers University, San Diego State University, Texas A&M, Ukrainian-American Birth Defects Program, University of Texas at Austin, University College London, UC Davis, UCLA, UC San Diego, Town University of Cape Town, University of Minnesota, University of New Mexico, University of North Carolina at Chapel Hill, and University of Southern California.

Survey Design/Methodology:

Available neurobehavior, dysmorphology demographics, 3D facial imaging, genetics, and

brain volume data for cross-sectional and longitudinal analyses were collected across one to six sites over three study phases. Four groups of children are included: children with prenatal alcohol exposure with or without a diagnosis of FAS; nonexposed typically developing controls; and two contrast groups of nonexposed children with other developmental conditions, including low IQ scores and the ADHD group.

Dysmorphology data were collected from all phases, and some people have been examined longitudinally. In Phase I (2003–2007), data were collected from children between the ages of 5 and 18 and from mothers and their babies in a prospective study of prenatal alcohol exposure. In Phase II (2007–2012), data were collected from children between the ages of 8 and 16, from 350 moderately to heavily alcohol-exposed pregnant women, and from approximately 350 low or unexposed pregnant women. Pregnancies were followed with serial ultrasounds, and maternal blood samples were collected and analyzed for various nutrients and markers of oxidative stress and inflammation. Live-born infants were given a dysmorphology exam. In Phase III (2012–2017), data were collected from children ages 5–7 and ages 10–16 under two age-based protocols. Data from maternal blood samples were collected from pregnant women, and fetal data were collected via ultrasound from mothers and babies. Collection of blood, urine, and cheek cell samples were also included.

Sample Characteristics:

Subject recruitment varied between the sites, and subject population was racially and ethnically diverse.

Alcohol Variables:

Detailed maternal history and patterns of alcohol use, drinking volume, drinking frequency, use and intensity of use of alcohol around pregnancy, alcohol use disorder, social influences on drinking, and paternal patterns of alcohol use were assessed.

Other Variables:

Included is information on the birth and alcohol exposure; dysmorphology data, including information on height, weight, head circumference, and other physical measures; 3D facial imaging; neurobehavioral examination data; psychological symptomatology; genetics; MRI brain volume; and cytokine data.

Collaborative Study on the Genetics of Alcoholism (COGA)—1991–2016, Ongoing

Sponsoring Agency:

National Institute on Alcohol Abuse and Alcoholism, National Institute on Drug Abuse, U.S. Department of Health and Human Services

Contact:

Washington University School of Medicine
Department of Psychiatry
660 South Euclid Avenue, Campus Box 8134
Saint Louis, MO 63110-1093

Sue Winkeler
Telephone: 314-286-2569 or Fax: 314-286-2577
winkeler@wustl.edu
<https://cogaproject.org/contact-information>

Availability:

Instructions for access to clinical data, genetic analysis, epidemiologic data, and biomaterials by several mechanisms subject to NIAAA approval are available at <https://cogastudy.org/access-coga-data#overlay-context>. COGA instruments are available for download from <https://cogastudy.org/coga-instruments>.

Overview:

COGA was funded in 1989 to identify the specific genes that can influence a person's likelihood of developing alcoholism. COGA consists of three inter-related scientific projects, supported by three cores, that work together to achieve the overarching goal of understanding the role of genes, functional networks, and neurobiological and environmental factors on risk and resilience. Extensive clinical, neuropsychological, electrophysiological, biochemical, and genetic data are collected to characterize the familial transmission of alcoholism and related phenotypes, and to identify susceptibility genes using genetic linkage. Researchers also established a repository of cell lines from respondents to serve as a permanent source of DNA for genetic studies.

Survey Design/Methodology:

Interviewing and testing of COGA families are conducted at 11 sites: SUNY Downstate Medical Center, University of Connecticut School of Medicine, Indiana University School of Medicine, University of Iowa, University of California San Diego, Howard University, Rutgers University, Icahn School of Medicine at Mt. Sinai, Virginia Commonwealth University, University of Pennsylvania, and Washington University School of Medicine. Eligible families are found using recruited patients currently in a psychiatric inpatient or outpatient program for alcohol and/or chemical dependency. Data collected from the families include blood biochemistry, psychological test performance, genetic analysis data consisting of marker genotypes, brain electrophysiological data,

lymphoblastoid cell lines, and DNA. The Semi-Structured Assessment for the Genetics of Alcoholism (SSAGA) developed for COGA is a polydiagnostic psychiatric interview that assesses substance use and mental health disorders. Companion instruments for children (ages 7–12), adolescents (ages 13–17), and parents (about their children) are also available. The Family History Assessment Module assesses *DSM-III-R* diagnoses for all family members.

Sample Characteristics:

Over the span of the COGA project, data have been collected in the following phases. Phase I: 1991–1998 (10,658 children [ages 7+], adolescents, and adults from 2,246 families at six data collection sites across the country); subjects completed interviews and questionnaires, underwent EEG testing, and provided DNA. Phase II: 1995–2005 (11,301 children, adolescents, and adults, including returning Phase I subjects, new participants from Phase I families, and Howard University subjects added to increase minority representation); subjects were given an updated Phase I assessment battery and cognitive tests. Phase III: 2003–2005 (474 children, adolescents, and adults drawn from Phases I and II). Prospective study: 2004–2019 (subjects ages 12–22 from original families evaluated every two years). One-Year Pilot Study of Older Adults: 2016–2017 (2,174 subjects ages 50+ years from Phases I and II who had not been seen for an average of 23 years and 706 subjects located and interviewed by telephone about physical and mental health and alcohol use/problems). Lifespan Phase: 2019–2024 (prospective

subjects ages 30+. Phase I/II subjects ages 50+); COVID-19 (telephone interviews conducted March–July 2020; mailed and web-based questionnaires).

Since 1991, COGA investigators have collected data on more than 2,255 extended families containing members who are affected by alcoholism. More than 17,702 people are represented in the database.

Alcohol Variables:

Alcohol dependence is measured using DSM-III-R and DSM-IV and at least one of the following diagnostic systems: Feighner

Research Diagnostic Criteria and ICD-10. Information on alcohol misuse and age of onset are also collected.

Other Variables:

The SSAGA covers demographic and medical history information, including tobacco use, marijuana and drug use, suicide attempts, anorexia, bulimia, adult attention-deficit/hyperactivity disorder, depression, mania, dysthymia, antisocial personality disorder, posttraumatic stress disorder, social phobia, and obsessive-compulsive disorder.

Coronary Artery Risk Development in Young Adults (CARDIA)—1985–1986, 1987–1988, 1990–1991, 1992–1993, 1995–1996, 2000–2001, 2005–2006, 2010–2011, and 2015–2016

Sponsoring Agency:

National Heart, Lung, and Blood Institute (NHLBI) of the National Institutes of Health, U.S. Department of Health and Human Services

Contact:

National Heart, Lung, and Blood Institute
6705 Rockledge Dr (RK1) Room 302-H
Mail Stop 7936
Bethesda MD 20892-7936
<https://biolincc.nhlbi.nih.gov/studies/cardia/?q=cardia>

Jared Reis, Ph.D.
Telephone: 301-435-1291 or Fax: 301-480-1455
Jared.Reis@nih.hhs.gov
CARDIAdataquestions@dopm.uab.edu

Availability:

NHLBI grants free limited-access to CARDIA data sets to researchers after they complete and submit an online application; at: <http://www.cardia.dopm.uab.edu/> or through the NHLBI Data Repository: <https://biolincc.nhlbi.nih.gov/studies/cardia/>. Study questionnaires and other documentation are available at <https://www.cardia.dopm.uab.edu/scientific-resources-landing-page2>.

Overview:

CARDIA is a prospective epidemiologic study designed to assess antecedents of cardiovascular disease (CVD) risk factors in young adults and to identify health behaviors predictive of levels and changes in these risk factors. Participants completed questionnaires and physical examinations and were followed for nine waves from 1985–1986 through 2015–2016. Repeat examinations were given to determine prevalence and changes in CVD risk factors.

Survey Design/Methodology:

Recruitment strategies were designed to achieve a representative sample of the underlying U.S. Black and White populations from four communities in diverse geographic regions. Sampling was by random-digit dialing or in-person contact with households in selected areas and census tracts in Birmingham, Chicago, and Minneapolis. Participants at the Oakland center were randomly selected from among subscribers of the Kaiser Permanente Medical Care Program living in specific residential areas around Oakland. Within each field center, attempts were made to achieve approximately equal numbers balanced on age, sex, self-reported race, and education status. Contact was maintained with participants every 6 months, with annual interim medical history

ascertainment; repeated assessment of genetic, psychosocial, neighborhood, environmental, lifestyle, and behavioral factors; prescription, recreational, and illicit drug use; and other information collected.

Sample Characteristics:

People ages 18–31 (n = 5,115) were recruited from the four community populations. These participants were asked to participate in follow-up examinations during 1987–1988 (Year 2), 1990–1991 (Year 5), 1992–1993 (Year 7), 1995–1996 (Year 10), 2000–2001 (Year 15), 2005–2006 (Year 20), 2010–2011 (Year 25), and 2015–2016 (Year 30). A majority of the group has been examined at each of the follow-up examinations (91 percent, 86 percent, 81 percent, 79 percent, 74 percent, 72 percent, 72 percent, and 71 percent, respectively). In Year 30, the sample consisted of 3,358 people that included those who were 47.8 percent Black, 43.0 percent male, and a mean age of 55.1 years.

Alcohol Variables:

At each examination, participants were asked about any past year drinking as well as follow-up questions to assess the number of drinks of wine, beer, and liquor they typically consumed in a week and their attempts to stop drinking. These data have been used to determine patterns of alcohol use and estimate alcohol

intake, and they may allow assessment of cumulative lifetime exposure.

Other Variables:

Data have been collected on a variety of factors believed to be related to heart disease in later life. These include conditions with clear links to heart disease, such as blood pressure, levels of cholesterol and other lipids, and blood glucose levels. Data have also been collected on physical measurements such as weight and skinfold fat; lifestyle factors such as substance use (tobacco, marijuana, opioids, and other nonmedical use of drugs), diet, and exercise patterns; behavioral and psychological variables; medical and family history; and other measures (e.g., insulin levels). In addition, echocardiography was

performed during Years 5, 10, 25, and 30; cardiac-computed tomography during Years 15, 20, and 25; carotid ultrasound during Year 20; measures of cognitive function at Years 25 and 30; and a brain magnetic resonance image in a subset of participants at Years 25 and 30. At Year 35, data collection will include first-time assessments of physical functioning (e.g., chair stand, grip strength), hearing, vestibular dysfunction, and dedicated lung and abdominal imaging as well as repeat measures of biological risk factors, sleep and accelerometry, cognitive function and brain imaging, dietary quality, and other measures. containing members who are affected by alcoholism. More than 17,702 people are represented in the database.

Drug Abuse Warning Network (DAWN)—1993–2003, 2004–2011, 2018–2021, Annually**Sponsoring Agency:**

Substance Abuse and Mental Health Services Administration (SAMHSA), U.S. Department of Health and Human Services

Contact:

Center for Behavioral Health Statistics and Quality
SAMHSA

5600 Fishers Lane, Parklawn Building

Rockville, MD 20857

Telephone: 240-276-1250 or Fax: 240-276-1260

CBHSQRequest@samhsa.hhs.gov

<https://www.samhsa.gov/data/data-we-collect/dawn-drug-abuse-warning-network>

Availability:

Public use data are available for download from <https://www.icpsr.umich.edu/icpsrweb/ICPSR/series/97>. Online reports and analysis can be found on the website.

Overview:

DAWN is a public health surveillance system that reports on drug-related visits to hospital emergency departments (EDs) and is used to monitor trends in drug use, identify emerging substances and drug combinations, assess health hazards associated with drug use and misuse, and estimate the impact of drug misuse on the nation's health care system. DAWN's target sample frame consists of all non-federal, short-stay, and general medical and surgical hospitals in the U.S. that have one or more EDs open 24 hours a day. DAWN captures both ED visits that are directly caused by drugs and those in which drugs are a contributing factor but not the direct cause of the ED visit. These criteria encompass all types of drug-related events, including misuse and accidental ingestion and adverse reactions. Because alcohol is considered an illicit drug for minors, alcohol misuse without the involvement of other drugs is considered a drug-related ED visit for patients under age 21. DAWN was redesigned in 2004, including changes that created a disruption in trends from the prior period. DAWN was reestablished in 2018 after a hiatus beginning in 2012 and redesigned for improved timeliness of data; data available at more frequent intervals; and data for a wider range of geographic types, including urban, suburban, and rural areas. DAWN now functions as a smaller-scale sentinel surveillance system, or an "early warning" system, focusing on detecting outbreaks—that is, sudden increases in ED visits for specific drugs; identifying new and novel psychoactive substances; monitoring the magnitude of the

health effects from substance use (as reflected in ED visits); and documenting the geographic, temporal, and demographic distribution of the problem.

Survey Design/Methodology:

DAWN relies on a longitudinal probability sample of hospitals located throughout the U.S., including Alaska and Hawaii. To be eligible for selection into the DAWN sample, a hospital must be a non-federal, short-stay, general medical and surgical hospital located in the U.S., with at least one 24-hour ED. Within each hospital, 50 to 100 percent of the days of the month are systematically selected, and a census of ED visits is selected for review for these days. DAWN cases are identified through the review of ED medical records in participating hospitals.

Sample Characteristics:

An estimated 141,529 ED visit cases were submitted in the redesigned 2021 DAWN, extrapolated to include an estimate of 2,942,609 alcohol-related ED visits. For 2011, 229,211 submitted cases were extrapolated to an estimate of 5,067,374 total drug-related ED visits. Considering the margin of error, this estimate may range from 4,616,753 to 5,517,995 drug-related ED visits out of the approximately 126 million total ED visits estimated for the universe of DAWN-eligible hospitals. Out of about 5.1 million drug-related ED visits, 2.5 million were considered to involve drug misuse, with the balance primarily involving adverse reactions and accidental ingestions.

Alcohol Variables:

Alcohol involvement is indicated for patients of all ages if it occurs with another drug. Because alcohol is considered an illicit drug for minors, ED visits involving “alcohol only” are included in DAWN for patients under age 21.

Other Variables:

Demographic variables include age (categorized), race/ethnicity, and sex. Also included are variables on substance use, including opioids, marijuana, and up to 22 other drugs among a comprehensive set of prescription and nonprescription drugs; the type of visit; and patient disposition.

Fatality Analysis Reporting System (FARS)—1975–2020, Annually

Sponsoring Agency:

National Highway Traffic Safety Administration (NHTSA), U.S. Department of Transportation

Contact:

National Center for Statistics and Analysis
NHTSA
1200 New Jersey Avenue, SE, West Building
Washington, DC 20590
NCSAWeb@nhtsa.dot.gov
<https://www.nhtsa.gov/research-data/fatality-analysis-reporting-system-fars>

Availability:

Data can be downloaded from <https://www.nhtsa.gov/file-downloads?p=nhtsa/downloads/FARS/> or <https://www.nhtsa.gov/file-downloads>. Online reports can be found here: <https://www-fars.nhtsa.dot.gov>.

Overview:

FARS is designed to assist the traffic safety community in identifying traffic safety problems (including drinking and driving), developing and implementing vehicle and driver countermeasures, and evaluating motor vehicle safety standards and highway safety initiatives. FARS gathers detailed data on all fatal traffic crashes each year within the 50 states, the District of Columbia, and Puerto Rico and has been in operation since 1975.

Survey Design/Methodology:

FARS is a census of all fatal traffic crashes. To be included, a crash must involve at least one motor vehicle moving on a roadway customarily open to the public and must result in the death of a person within 30 days of the crash. Each case has more than 100 data elements that characterize the crash and are coded at four levels: the accident, the vehicle, the driver, and the person(s) involved. Data sources may include police crash reports, state vehicle registration files, state driver licensing files, state highway department files, vital statistics documents, death certificates, coroner reports, hospital reports, and emergency medical services reports. The specific data elements may be modified slightly over the years.

Sample Characteristics:

The total number of FARS cases varies from year to year. In 2018, FARS reported 33,919 fatal traffic crashes that resulted in 36,835 deaths.

Alcohol Variables:

Alcohol variables include the alcohol test type and results, police-reported alcohol involvement, and the method of alcohol determination by police. The driver is considered a drunk driver if the blood alcohol concentration (BAC) is positive, or if the police reported alcohol involvement. Since 1984, NHTSA has used statistical methods to estimate BAC values for drivers with unknown BAC levels. The imputed BAC data are provided in separate data files.

Other Variables:

Other variables include reported use of, or intoxication from opioids, marijuana, and other drugs; age, sex, role (driver, passenger, nonoccupant) for all people in the traffic crash; injury severity, time, and date of the crash; number of vehicles involved; vehicle make and model; speed limit; road and atmospheric conditions; violations charged; and previous convictions of traffic violations for all drivers.

Healthcare Cost and Utilization Project (HCUP) Kids' Inpatient Database (KID)— 1997–2012, Triennially, and 2016 and 2021

Sponsoring Agency:

Agency for Healthcare Research and Quality (AHRQ), U.S. Department of Health and Human Services

Contact:

AHRQ

Office of Communications

5600 Fishers Lane, 7th Floor

Rockville, MD 20857

hcup@ahrq.gov

<https://www.hcup-us.ahrq.gov/db/nation/kid/kiddbdocumentation.jsp>

Availability:

Summary statistics through 2016 are available at <https://www.hcup-us.ahrq.gov/db/nation/kid/kidsummarystats.jsp>. Data distributed via digital download are available for purchase from the HCUP Central Distributor, c/o IBM Watson Health, 5425 Hollister Ave., Suite 140, Santa Barbara, CA 93111, Phone: (866) 290-HCUP (4287), Fax: (805) 979-3787, Email: hcup@ahrq.gov.

Overview:

HCUP is a federal-state-industry partnership in health care data collection. It includes patient data from all payer sources. HCUP's objectives are to (1) create a source for national, state, and all-payer health care data; (2) produce a set of tools and products to facilitate the use of HCUP and other administrative data; (3) enrich a collaborative partnership with statewide data organizations to increase the quality and use of health care data; and (4) conduct and translate research to inform decision making and improve health care delivery. HCUP data allow for comparative studies of health care services and the use and cost of hospital care, including effects of market forces on hospitals and the care they provide, variations in medical practice, effectiveness of medical technology and treatments, and use of services by special populations. KID is an all-payer pediatric inpatient care database in the U.S., yielding national estimates of children's hospital inpatient stays.

Survey Design/Methodology:

KID is a sample of pediatric discharges from all community, nonrehabilitation hospitals in states participating in HCUP. Its unit of observation is an inpatient stay record. The core file contains data elements for linkage, patient demographics, clinical information, and payment information. The hospital-level file

contains one observation for each hospital included in the KID and contains variance estimation data elements, linkage data elements, and data elements that describe basic hospital characteristics.

Sample Characteristics:

The numbers of states participating in the respective surveys were: 22 (1997), 27 (2000), 36 (2003), 38 (2006), 44 (2009, 2012), and 46 and the District of Columbia (2016). In 2016, all states except for Alabama, Delaware, Idaho, and New Hampshire participated in the KID. The numbers of pediatric discharges included were: 2,521 (1997), 2,784 (2000), 3,438 (2003), 3,739 (2006), 4,121 (2009), 4,179 (2012), and 4,200 (2016).

Alcohol Variables:

Alcohol use-related diagnoses are coded in ICD-9 and ICD-10 fields, and an alcohol use/misuse comorbidity measure is also included in the data set.

Other Variables:

KID includes other key variables such as principal diagnosis, any listed diagnosis with ICD-9-CM codes, ICD-10-CM codes since 2016 (including other substance use disorders), and a drug use/misuse comorbidity measure. Other variables include hospital characteristics, diagnostic-related groups, severity of illness measures, and comorbidity measures.

Healthcare Cost and Utilization Project (HCUP) Nationwide Ambulatory Surgery Sample (NASS)—2016–2019, Annually

Sponsoring Agency:

Agency for Healthcare Research and Quality (AHRQ), U.S. Department of Health and Human Services

Contact:

AHRQ

Office of Communications
5600 Fishers Lane, 7th Floor
Rockville, MD 20857

hcup@ahrq.gov

<https://www.hcup-us.ahrq.gov/nassoverview.jsp>

Availability:

Summary statistics through 2018 are available at:

<https://www.hcup-us.ahrq.gov/db/nation/nass/nasssummstats.jsp>. Data for 2016–2018 are available for purchase from HCUP: HCUP Central Distributor, c/o IBM Watson Health, 5425 Hollister Ave., Suite 140, Santa Barbara, CA 93111, Phone: (866) 290-HCUP (4287), Fax: (805) 979-3787, Email:

hcup@ahrq.gov

Overview:

HCUP is a federal-state-industry partnership in health care data collection. It includes inpatient data from all payer sources. HCUP's objectives are to (1) create a source for national, state, and all-payer health care data; (2) produce a set of tools and products to facilitate the use of HCUP and other administrative data; (3) enrich a collaborative partnership with statewide data organizations to increase the quality and use of health care data; and (4) conduct and translate research to inform decision making and improve health care delivery. HCUP data allow for comparative studies of health care services and the use and cost of hospital care, including the effects of market forces on hospitals and the care they provide, variations in medical practice, the effectiveness of medical technology and treatments, and use of services by special populations. NASS was created to enable analyses of selected ambulatory surgery utilization patterns and to support public health professionals, administrators, policymakers, and clinicians in their decision making regarding ambulatory surgery care.

Survey Design/Methodology:

NASS is designed to be representative of U.S. hospital-owned facilities that perform ambulatory surgeries. NASS contains clinical and resource-use information that is included in a typical hospital-owned facility record abstract, including patient characteristics, clinical diagnostic and surgical procedure codes,

disposition of patients, total charges, expected source of payment, and facility characteristics.

Sample Characteristics:

NASS is sampled from the HCUP State Ambulatory Surgery and Services Databases (SASD), which include various types of outpatient services, such as observation stays, lithotripsy, radiation therapy, imaging, chemotherapy, and labor and delivery. It contains information from 7.7 million ambulatory surgery encounters at 2,699 hospital-owned facilities that approximate an estimated 62 percent stratified sample of U.S. hospital-owned facilities performing ambulatory surgeries. The specific types of ambulatory surgeries and outpatient services included in each SASD vary by state and data year. SASD do not include ambulatory surgery encounters subsequently admitted to the same hospital for inpatient care. The number of states, including the District of Columbia, per year were 32 in 2016, 32 in 2017, and 34 in 2018. NASS states with the District of Columbia accounted for 82 percent of the U.S. population in 2018, an estimated 62 percent of hospital-owned facilities performing ambulatory surgeries, and an estimated 72 percent of ambulatory surgery encounters.

Alcohol Variables:

NASS contains alcohol-related diagnoses coded according to ICD-10-CM and may be analyzed by geographic region, hospital ownership,

urban/rural location, and quality-of-care outcomes.

Other Variables:

Substance use-related diagnosis variables other than alcohol are coded according to ICD-10-CM.

NASS includes other variables, including patient demographics, expected payment source, total charges, disposition of the patient, and hospital characteristics.

Healthcare Cost and Utilization Project (HCUP) Nationwide Emergency Department Sample (NEDS)—2006–2019, Annually

Sponsoring Agency:

Agency for Healthcare Research and Quality (AHRQ), U.S. Department of Health and Human Services

Contact:

AHRQ

Office of Communications
5600 Fishers Lane, 7th Floor
Rockville, MD 20857

hcup@ahrq.gov

<https://www.ahrq.gov/data/hcup/index.html>

<https://www.hcup-us.ahrq.gov/overview.jsp>

Availability:

Summary statistics through 2019 are available at <http://www.hcup-us.ahrq.gov/db/nation/neds/nedssummstats.jsp>. Data for 2006–2019 are available for purchase from the HCUP Central Distributor, c/o IBM Watson Health, 5425 Hollister Ave., Suite 140, Santa Barbara, CA 93111, Phone: (866) 290-HCUP (4287), Fax: (805) 979-3787, Email: hcup@ahrq.gov

Overview:

HCUP is a federal-state-industry partnership in health care data collection. It includes patient data from all payer sources. HCUP's objectives are to (1) create a source for national, state, and all-payer health care data; (2) produce a set of tools and products to facilitate the use of HCUP and other administrative data; (3) enrich a collaborative partnership with statewide data organizations to increase the quality and use of health care data; and (4) conduct and translate research to inform decision making and improve health care delivery. HCUP data allow for comparative studies of health care services and the use and cost of hospital care, including effects of market forces on hospitals and the care they provide, variations in medical practice, effectiveness of medical technology and treatments, and use of services by special populations. NEDS, a part of HCUP, is a database containing patient-level information on emergency department (ED) visits across the country.

Survey Design/Methodology:

NEDS was constructed using the HCUP State Emergency Department Databases (SEDD) and the State Inpatient Databases (SID). SEDD capture discharge information on ED visits that do not result in a hospital admission (i.e., treat-and-release visits and transfers to another hospital). SID contain information on patients initially seen in the ED and then admitted to the same hospital. NEDS uses a stratified probability sample of U.S. hospital-based EDs.

NEDS includes all visits within the sample of selected EDs.

Sample Characteristics:

As the largest U.S. publicly available all-payer ED visits database, NEDS contains data on ED visits at more than 940 hospitals, approximating a 20-percent sample of U.S. hospital-based EDs. The number of states involved per year is: 2006 (24 states), 2007 (27 states), 2008 (28 states), 2009 (29 states), 2010 (28 states), 2011–2013 (30 states), 2014 (33 states and the District of Columbia), and 2019 (40 states and the District of Columbia).

Alcohol Variables:

NEDS contains fields for up to 35 ED visit diagnoses, including alcohol-related conditions starting in 2017 (30 diagnoses for years 2014–2016 and 15 ED visit diagnoses prior to 2014). ICD-9-CM codes are included through the first three quarters of 2015, and ICD-10-CM/PCS are included thereafter.

Other Variables:

NEDS includes other variables such as principal diagnosis and any listed diagnosis (including other substance use disorders); principal procedure, any listed procedure, and number of procedures; disposition of the patient at ED discharge; diagnosis-related group in effect on discharge; age, race, and sex; death during hospitalization; length of stay; primary and secondary payer; and income. In 2009, NEDS added a series of data elements that identified injuries by severity, mechanism, and intent.

Healthcare Cost and Utilization Project (HCUP) National (Nationwide) Inpatient Sample (NIS)—1988–2019, Annually

Sponsoring Agency:

Agency for Healthcare Research and Quality (AHRQ), U.S. Department of Health and Human Services

Contact:

AHRQ

Office of Communications

5600 Fishers Lane, 7th Floor

Rockville, MD 20857

hcup@ahrq.gov

<https://www.ahrq.gov/data/hcup/index.html>

<https://www.hcup-us.ahrq.gov/nisoverview.jsp>

Availability:

Summary statistics through 2019 are available at <http://www.hcup-us.ahrq.gov/db/nation/nis/nisummstats.jsp>. Data for 1988–2019 are available for purchase from HCUP Central Distributor, c/o IBM Watson Health, 5425 Hollister Ave., Suite 140, Santa Barbara, CA 93111, Phone: (866) 290-HCUP (4287), Fax: (805) 979-3787, Email: hcup@ahrq.gov

Overview:

HCUP is a federal-state-industry partnership in health care data collection. It includes inpatient data from all payer sources. HCUP's objectives are to (1) create a source for national, state, and all-payer health care data; (2) produce a set of tools and products to facilitate the use of HCUP and other administrative data; (3) enrich a collaborative partnership with statewide data organizations to increase the quality and use of health care data; and (4) conduct and translate research to inform decision making and improve health care delivery. HCUP data allow for comparative studies of health care services and the use and cost of hospital care, including the effects of market forces on hospitals and the care they provide, variations in medical practice, the effectiveness of medical technology and treatments, and use of services by special populations. The National Inpatient Sample (NIS—formerly the Nationwide Inpatient Sample), part of HCUP, is a database containing patient-level information on inpatient hospital stays.

Survey Design/Methodology:

Before 2012, NIS was a sample of hospitals, and all discharges from the hospitals were retained. In 2012, NIS was redesigned as a sample of discharges from all hospitals participating in HCUP. NIS examines hospital inpatient stays derived from billing data submitted by hospitals to statewide data organizations across the U.S. Inpatient stay records include clinical and

resource use information typically available from discharge abstracts. Discharge weights are provided for national estimates. As of 2012, NIS defines hospitals and discharges using the definitions supplied by the statewide data organizations that contribute to HCUP rather than the definitions used by the AHA Annual Survey. Changes starting in 2012 also include eliminating state and hospital identifiers, meaning that hospital linkages can no longer be performed.

Sample Characteristics:

NIS is sampled from the State Inpatient Databases (SID), which include all inpatient data currently contributed to HCUP. NIS includes all patients, regardless of health insurance status. NIS is drawn from all states participating in HCUP, covering more than 96 percent of the U.S. population and approximating a 20 percent stratified sample of discharges from U.S. community hospitals. Data include more than 7 million inpatient stays. Data releases and the number of states involved are: Release 1 Data: 1988–92 (8 states in 1988, 11 in 1989–92). Releases 2–23 Data: 1993 (17 states), 1994 (17 states), 1995 (19 states), 1996 (19 states), 1997 (22 states), 1998 (22 states), 1999 (24 states), 2000 (28 states), 2001 (33 states), 2002 (35 states), 2003 (37 states), 2004 (37 states), 2005 (37 states), 2006 (38 states), 2007 (40 states), 2008 (42 states), 2009 (44 states), 2010 (45 states), 2011 (46 states), 2012–2013 (44 states), 2014 (45 states), and 2019 (48 states and the District of Columbia).

Alcohol Variables:

NIS contains alcohol-related diagnoses (coded according to ICD-10-CM since the fourth quarter of 2016 and ICD-9-CM before then) that may be analyzed by geographic region, hospital ownership, urban/rural location, and quality-of-care outcomes.

Other Variables:

NIS includes other key variables such as principal diagnosis and any listed diagnosis (including other substance use disorders); principal procedure and any listed procedure; diagnosis-related group in effect on discharge; age, race, and sex; death during hospitalization; length of stay; primary and secondary payer; and income.

Healthcare Cost and Utilization Project (HCUP) National Readmissions Database (NRD)—2010–2018, Annually

Sponsoring Agency:

Agency for Healthcare Research and Quality (AHRQ), U.S. Department of Health and Human Services

Contact:

AHRQ

Office of Communications

5600 Fishers Lane, 7th Floor

Rockville, MD 20857

hcup@ahrq.gov

<https://www.hcup-us.ahrq.gov/nassoverview.jsp>

Availability:

NRD data are available for purchase from: HCUP Central Distributor, c/o IBM Watson Health, 5425 Hollister Ave., Suite 140, Santa Barbara, CA 93111, Phone: (866) 290-HCUP (4287), Fax: (805) 979-3787, Email: hcup@ahrq.gov

Overview:

HCUP is a federal-state-industry partnership in health care data collection. It includes inpatient data from all payer sources. HCUP's objectives are to (1) create a source for national, state, and all-payer health care data; (2) produce a set of tools and products to facilitate the use of HCUP and other administrative data; (3) enrich a collaborative partnership with statewide data organizations to increase the quality and use of health care data; and (4) conduct and translate research to inform decision making and improve health care delivery. HCUP data allow for comparative studies of health care services and the use and cost of hospital care, including the effects of market forces on hospitals and the care they provide, variations in medical practice, the effectiveness of medical technology and treatments, and use of services by special populations. NRD is drawn from the HCUP State Inpatient Databases (SID) and was created to enable analyses of national readmission rates.

Survey Design/Methodology:

The 2018 NRD was constructed from 28 states with reliable, verified patient linkage numbers in SID that could be used to track the patient across hospitals within a state. NRD is an annual file containing inpatient records for patients discharged in a calendar year. The files

included patients admitted in the prior year and discharged in the current year and excluded patients admitted to a hospital in the current year but discharged in the next year.

Sample Characteristics:

The 28 states are geographically dispersed and account for 59.7 percent of the total U.S. resident population and 58.7 percent of all U.S. hospitalizations.

Alcohol Variables:

Alcohol use-related diagnoses are coded in ICD-9 and ICD-10 fields, and an alcohol use/misuse comorbidity measure is also included in the data set.

Other Variables:

NRD comprises more than 100 clinical and nonclinical variables for each hospital stay, with ICD-9-CM codes and ICD-10-CM/PCs since 2016. These include substance use disorders other than alcohol; a drug use/misuse comorbidity measure; other diagnoses, procedures; external cause of injury codes; patient demographics (sex, age, median household income quartile, and urban/rural location of the patient's residence); and expected payment source (Medicare, Medicaid, private insurance, self-pay, those billed as "no charge," and other insurance types).

Health Information National Trends Survey (HINTS)—2003, 2005, 2007, 2009, 2011–2015, and 2017–2020, Annually

Sponsoring Agency:

National Cancer Institute of the National Institutes of Health, U.S. Department of Health and Human Services

Contact:

Division of Cancer Control and Population Sciences
National Cancer Institute
9609 Medical Center Drive, MSC 9671
Bethesda, MD 20892-9671
NCIhints@mail.nih.gov
<https://hints.cancer.gov/about-hints/contact-us.aspx>

Availability:

All public-use HINTS data sets and documentation are free to download at <https://hints.cancer.gov/data/download-data.aspx>. Access to restricted-use data sets that contain geocodes and suppressed variables can be requested by application at <https://hints.cancer.gov/data/restricted-data.aspx>. Summary statistics by survey iteration and survey item can be obtained using the online tool at <https://hints.cancer.gov/view-questions-topics/all-hints-questions.aspx>

Overview:

HINTS began in 2003 and has been administered annually since 2011; it is a biennial, cross-sectional survey of a nationally representative sample of the U.S. civilian, noninstitutionalized, adult population used to assess the impact of the health information environment. Specifically, HINTS measures how people access and use health information, how people use information technology to manage health and health information, and the degree to which people are engaged in healthy behaviors. Several items have a specific focus on cancer prevention and control. The survey is designed to provide updates on changing patterns, needs, and information opportunities in health; identify changing communications trends and practices; assess cancer information access and usage; provide information about how cancer risks are perceived; and offer a testbed to researchers to test new theories in health communication. HINTS was developed by the Health Communication and Informatics Research Branch of the Division of Cancer Control and Population Sciences of the National Cancer Institute. There are 14 iterations of HINTS. As such, researchers can examine items and constructs that are common to all iterations to measure trends over time.

Survey Design/Methodology:

HINTS is conducted on a weighted probability sample of adults ages 18 and older, with an oversampling of Hispanic and African American households at baseline, including English speakers and Spanish speakers collected through translated versions. HINTS 1–3 were administered by landline telephone using a random-digit-dial sample frame. HINTS 3 was also conducted using self-administered mail questionnaires, as were all iterations of HINTS 4 and HINTS 5. HINTS data provide nationally representative estimates, but variables in the data set allow researchers to compare rural vs. urban metropolitan statistical areas as well as four census regions and nine census divisions. In certain cases, data can be pooled across multiple iterations of HINTS so that the sample size may be sufficiently large to perform a state-based investigation.

Sample Characteristics:

The 14 survey iterations and their respective sample sizes were: HINTS 1 (2003) ($n = 6,369$); HINTS 2 (2005) ($n = 5,586$); HINTS 3 (2007) ($n = 7,674$); HINTS Puerto Rico (2009) ($n = 639$); HINTS 4 Cycle 1 (2011) ($n = 3,959$); HINTS 4 Cycle 2 (2012) ($n = 3,630$); HINTS 4 Cycle 3 (2013) ($n = 3,185$); HINTS 4 Cycle 4 (2014) ($n = 3,677$); HINTS-FDA (2015) ($n = 3,738$); HINTS-FDA Cycle 2 (2017) ($n = 1,736$); HINTS 5 Cycle 1 (2017) ($n = 3,285$); HINTS 5 Cycle 2 (2018) (n

= 3,504); HINTS 5 Cycle 3 (2019) ($n = 5,438$); and HINTS 5 Cycle 4 (2020) ($n = 3,865$). The HINTS 5 Cycle 4 (2020) population sample had the following characteristics: age: 18–34 years (25.5 percent), 35–49 years (24.8 percent), 50–64 years (27 percent), 65–74 years (11.6 percent), ≥ 75 years (8.4 percent); sexual orientation: heterosexual/straight (88.5 percent), homosexual/gay/lesbian (2.5 percent), bisexual (2.6 percent); race/ethnicity: non-Hispanic White (58.7 percent), Hispanic or Latino (15.7 percent), non-Hispanic Black (10.3 percent), non-Hispanic American Indian/Alaskan Native (0.5 percent), non-Hispanic Asian (4.8 percent), non-Hispanic Hawaiian/Pacific Islander (0.5 percent), and non-Hispanic multiracial (2.1 percent).

Alcohol Variables:

Alcohol use questions include current drinking frequency and intensity, perceptions of cancer

and other health risks from alcohol, and attitudes toward alcohol control policies.

Other Variables:

Substance use variables other than alcohol include use of tobacco products and electronic cigarettes. Other topics covered include: breast cancer, cancer communication, cancer perceptions and knowledge, caregiving, cervical cancer, clinical trials, colon cancer, demographics, food and medical products information, genetic testing, health communication, health services, health status, internet use, lung cancer, medical research and medical records, numeracy, nutrition and physical activity, palliative care, patient-provider communication, prostate cancer, risk perceptions, skin cancer, skin protection, and social networks.

National Alcohol Survey (NAS)—1964–1965, 1967, 1969, 1974, 1979, 1984, 1990, 1992, 1995–1996, 2000–2001, 2005, 2009–2010, and 2014–2015

Sponsoring Agency:

Alcohol Research Group, Public Health Institute, and funded by the National Institute on Alcohol Abuse and Alcoholism, U.S. Department of Health and Human Services

Contact:

Alcohol Research Group
Public Health Institute
6001 Shellmound St., Suite 450
Emeryville, CA 94608-1010
Telephone: 510-898-5800
info@arg.org
<http://arg.org/center/national-alcohol-surveys/>

Availability:

The N12 and earlier national data and documentation from NAS are available on request from <http://arg.org/nas-data-sets/>. Requests for N13 data will be evaluated on a case-by-case basis; contact study director Dr. Priscilla Martinez at pmartinez@arg.org.

Overview:

NAS is designed to assess trends in drinking practices and problems in the national population, including drinking patterns, attitudes, norms, treatment experiences, and adverse consequences. NAS also studies the effects of public policy on drinking practices (i.e., beverage warning labels).

Survey Design/Methodology:

NAS used a multistage-area probability sample of people ages 18 and older in households within the 48 contiguous states through 1995 (i.e., survey N9). Starting in 2000, NAS has used a random digit dialing sampling and computer-assisted telephone interviewing of adults in households in all 50 states and the District of Columbia. Blacks and Hispanics were oversampled in N7 and N9 and later NAS surveys. Special populations in various institutional settings, including detoxification centers, jails, clinics, emergency rooms, and welfare offices, were not included.

Sample Characteristics:

In 2014–2015 (Survey N13), the sample consisted of 7,071 respondents ages 18 and older, including

oversamples of Black Non-Hispanics ($n = 1,763$) and Hispanics ($n = 1,623$).

Alcohol Variables:

NAS data are collected on graduated frequencies and other measures of lifetime and current alcohol consumption; beverage type including beer, wine, and spirits; binge drinking; attempts to reduce drinking; attitudes/opinions on drinking levels in different drinking situations; treatment status; and drinking consequences. Drinking problems include alcohol dependence symptoms, life area harms, and tangible consequences such as employment repercussions, injury or health effects, and psychological/emotional distress.

Other Variables:

Demographic variables include age, race, sex, geographic region, education, income, and others. Other variables include use of marijuana, tobacco, and other substances; and attitudes and values concerning violence, injury, risk-taking behaviors, substance use, illegal behaviors, arrests, and convictions.

National Ambulatory Medical Care Survey (NAMCS)—1973–1992, 1993–2016, 2018, Annually

Sponsoring Agency:

National Center for Health Statistics (NCHS), Centers for Disease Control and Prevention

Contact:

Ambulatory and Hospital Care Statistics Branch
NCHS
3311 Toledo Road
Hyattsville, MD 20782
Telephone: 301-458-4600
ambcare@cdc.gov
<http://www.cdc.gov/nchs/ahcd.htm>

Availability:

Data files are available for download from
https://www.cdc.gov/nchs/ahcd/ahcd_questionnaires.htm#public-use.

Overview:

NAMCS is designed to provide reliable information about the provision and use of ambulatory medical care services in the U.S., including ambulatory care visits to physician offices, physician practices, and patient and visit characteristics.

Survey Design/Methodology:

NAMCS uses a multistage probability design involving probability samples of primary sampling units (PSUs), physician practices within PSUs, and patient visits within practices. First-stage samples include PSUs that are counties, groups of counties, county equivalents (such as parishes or independent cities), or towns and townships. Second-stage samples consist of a probability sample of practicing physicians contained in master files maintained by the American Medical Association and American Osteopathic Association. The physicians are office based, principally engaged in patient care activities; non-federally employed; and not in the specialties of anesthesiology, pathology, and radiology. All eligible physicians are stratified into 15 groups, and a sample is taken from their patient visits. The physician sample is divided into 52 random subsamples and assigned to 1 of the 52 weeks in the survey year. Random patient visit samples are selected by the physician during an assigned week. Physicians collect the actual data, aided by their office staff when possible. Survey design changes in 2012 made state-level estimates possible for the 34 most populous states and the 9 Census divisions. The 2013 and 2014 surveys used

state-based samples but included the 22 and 17 (plus Wisconsin) most populous states, respectively. In 2012, NAMCS switched from a paper-and-pencil mode of data collection to a computer-assisted method.

Sample Characteristics:

NAMCS sample sizes of patients vary from year to year, from a 100-percent sample for very small practices to a 10-percent sample for very large practices. During 2018, NAMCS collected a total of 9,953 patient record forms from 496 physicians, a sample reflecting 860 million U.S. office visits.

Alcohol Variables:

Alcohol use or alcohol-related conditions cited as a reason for the visit are coded when mentioned by the patient. In 2014, alcohol misuse or dependence were added to each provider's diagnosis for visit, and alcohol misuse screening and counseling were added to the services section.

Other Variables:

Patient variables include date of visit, age, sex, race, ethnicity, reason for visit (up to three), expected source(s) of payment, diagnostic screening services, and physician's diagnoses (up to three). Also included are referral and previous visit history, medication and nonmedication therapy (up to five medications), disposition and duration of visit, weight, geographic region (prior to 2018), and SMSA code. In 1997, pregnancy status, authorization requirements, HMO status, and the major reason for the patient visit were added to the NAMCS.

National Automotive Sampling System General Estimates System (NASS GES)—1988–2015, Crash Report Sampling System (CRSS)—2016–2020, Annually

Sponsoring Agency:

National Highway Traffic Safety Administration (NHTSA), U.S. Department of Transportation

Contact:

National Highway Traffic Safety Administration Headquarters

National Center for Statistics & Analysis

1200 New Jersey Avenue SE

West Building

Washington, DC 20590

Automated Data Request Line: 800-934-8517

NCSARequests@dot.gov, nhtsa.webmaster@dot.gov

<https://www.nhtsa.gov/national-automotive-sampling-system-nass/nass-general-estimates-system#nass-general-estimates-system-overview>

<https://www.nhtsa.gov/crash-data-systems/crash-report-sampling-system>

Availability:

NASS GES data can be downloaded from <https://www.nhtsa.gov/file-downloads?p=nhtsa/downloads/NASS/>

CRSS data can be downloaded from <https://www.nhtsa.gov/file-downloads?p=nhtsa/downloads/CRSS/>.

Customized queries can be constructed using an online query tool at <https://cdan.dot.gov/query>.

Overview:

NASS GES began in 1988 to support the development, implementation, and assessment of highway safety programs aimed at reducing the human and economic cost of motor vehicle traffic crashes. The CRSS system was built on NASS GES and replaced it since 2016. Its purpose is to create an estimate of safety-related vehicle crashes in the U.S.; identify highway safety problem areas; measure vehicle crash and safety trends; drive consumer information initiatives; and form the basis for cost and benefit analyses of highway safety initiatives, countermeasures, and regulations, population, miles driven, and number of crashes in the U.S.

Survey Design/Methodology:

CRSS's scope is vehicle crashes that involve at least one motor vehicle in transport on a traffic way that resulted in a fatality, serious injury, or property damage. CRSS data are derived from a statistically representative sample of police-reported crashes involving all types of motor vehicles, pedestrians, and cyclists and weighted to reflect the geography, divided into three stratification stages: (1) primary sampling unit (a county or group of counties), (2) police jurisdiction (PJ), and (3) police accident report (PAR) (filled out by police officers and reported to the state through the police jurisdiction). In this way, PJs are viewed as natural clusters of PARs.

Sample Characteristics:

CRSS uses a sample of PARs from an estimated 5-to-7 million annual police-reported crashes. In 2018, police crash reports were sampled and coded from 389 PJs at 60 selected sites across the U.S. There were an estimated 6,734,000 police-reported motor vehicle traffic crashes, resulting in 36,560 fatalities and 2,710,000 people injured. Among these crashes, less than 1 percent were fatal crashes, 28 percent were injury crashes, and 71 percent were property-damage-only crashes.

Alcohol Variables:

Alcohol use by anyone in the traffic crash is recorded based on police-reported alcohol involvement. The method of alcohol determination by police, the type of test used, and the test result are recorded. Alcohol use is imputed for people with an unknown value on this variable. Also included is a variable indicating violation(s) charged to the vehicle drivers, including driving under the influence of alcohol and/or drugs.

Other Variables:

Other key variables include presence of other drugs and test results, age, sex, time and date of occurrence, crash scenario and contributing factors, injury information, fatalities, property damage, vehicle make, and sample weights.

National COVID Cohort Collaborative (N3C) Data Enclave—2020–2022, Ongoing

Sponsoring Agency:

National Center for Advancing Translational Sciences (NCATS)

N3C is composed of members from the National Institutes of Health Clinical and Translational Science Awards Program and its Center for Data to Health, the IDeA Centers for Translational Research, the National Patient-Centered Clinical Research Network, the Observational Health Data Sciences and Informatics network, TriNetX, and the Accrual to Clinical Trials network.

Contact:

NCATS

6701 Democracy Boulevard

Bethesda MD 20892-4874

Telephone: 301-594-8966

NCATS_N3C@nih.gov

<https://covid.cd2h.org/>

Availability:

Find instructions for applying for access to the N3C at <https://ncats.nih.gov/n3c/about/applying-for-access>. Three tiers of access are available. Requirements for the least restrictive tier providing “synthetic” (i.e., computationally derived data from a limited data set) include: (1) a data use agreement; (2) completion of the NIH “Information Security, Counterintelligence, Privacy Awareness, Records Management Refresher, Emergency Preparedness Refresher” course; and (3) submission of a data use request through the N3C data enclave with a project title, a public research statement, a description of the research project plan, and the level of data intended to access and other information. A nonconfidential abstract allows genomic, radiology, pathology imaging, and other data to be analyzed in conjunction with the N3C clinical data.

Overview:

N3C was formed in April 2020 as a centralized resource of harmonized electronic health record (EHR) data from multiple health systems around the U.S. to accelerate understanding of SARS-CoV-2 (severe acute respiratory syndrome coronavirus 2) infection, the virus that causes COVID-19 disease. The N3C data enclave provides a diverse and nationally representative central repository of harmonized EHR data and represents a new model for collaborative data sharing and analytics. The N3C architecture, data set, and analytic environment form a platform for developing machine learning algorithms, statistical models, and clinical decision support tools. Analytic models can use time series, clinical, and laboratory information to predict progression, assess need and efficacy of clinical interventions, and predict long-term sequelae. The platform supports translational informatics in the form of knowledge graphs and related tools, mined and annotated literature, and clinical EHR data.

Survey Design/Methodology:

NCATS asks medical institutions and health care organizations to contribute a limited data set, pursuant to the requirements in the Health Insurance Portability and Accountability Act Privacy Rule. N3C systematically and regularly collects data derived from the EHRs of people who were tested for COVID-19 or who had related symptoms as well as data from people infected with pathogens that can support comparative studies, such as SARS 1, MERS, and H1N1. The data set includes information such as demographics, symptoms, lab test results, procedures, medications, medical conditions, and physical measurements.

The analytics platform or N3C enclave, hosted by a secure NCATS-controlled cloud environment, includes clinical data from patients who meet criteria in the N3C COVID-19 phenotype from sites across the U.S. dating back to January 2018. The full N3C data enclave includes EHR data of patients from partner sites who were tested for COVID-19 or

had related symptoms after January 1, 2020. For patients included in the N3C data enclave, encounters in the same source health system beginning on or after January 1, 2018 are also included to provide lookback data. N3C utilizes centrally maintained shared logic sets for common diagnostic and phenotype definitions.

Sample Characteristics:

The N3C cohort explorer provides up-to-date information about N3C data at <https://covid.cd2h.org/dashboard/cohort>. As of April 7, 2022, 296 institutions have executed data use agreements, and 72 sites had harmonized data included in the N3C data

enclave pertaining to 13.0 million people, 1.3 billion clinical observations, 6.9 billion lab results, 2.2 billion medical records, 663.4 million procedures, and 710.6 million visits.

Alcohol Variables:

ICD-9-CM diagnosis codes and other condition codes are used to identify alcohol-related symptoms morbidity.

Other Variables:

Other variables include patient demographic characteristics, medical and family history, diagnostic and therapeutic procedures and results, disposition, and other clinical information.

National Crime Victimization Survey (NCVS)—1973–2020, Annually

Sponsoring Agency:

Bureau of Justice Statistics (BJS), Office of Justice Programs, U.S. Department of Justice

Contact:

Bureau of Justice Statistics

810 Seventh Street, NW

Washington, DC 20531

Telephone: 202-307-0765

askbjs@usdoj.gov

<http://bjs.ojp.usdoj.gov/index.cfm?ty=dcdetail&iid=245>

Availability:

Data files for 1992–2019 are available for download from

<http://www.icpsr.umich.edu/icpsrweb/ICPSR/series/95>. Contact BJS for information about other NCVS data available on CD-ROM.

Overview:

NCVS—formerly the National Crime Survey—collects data on personal and household victimization in the U.S. The program has four primary objectives: to develop detailed information about the victims and consequences of crime, to estimate the numbers and types of crimes not reported to the police, to provide uniform measures of selected types of crimes, and to permit comparisons over time and types of areas. A School Crime Supplement was conducted in 1989, 1995, and 1999–2013, biennially, studying students ages 12 to 19 (ages 12 to 18 since 1999) in school programs leading toward diplomas. NCVS was redesigned in 1992 to improve data on sexual assaults and domestic violence and to improve recall ability.

Survey Design/Methodology:

NCVS is an ongoing national probability survey of residential addresses in selected U.S. cities using a stratified multistage cluster sample. Data are collected quarterly, and six quarters comprise an annual file (four in the current year and the first two quarters of the following year). NCVS data are collected by telephone and through in-person interviews. Several methodological changes were implemented in 2006 that affected the victimization rate estimates for that year. These effects were

reversed in 2007, suggesting that the 2006 findings represent a temporary anomaly in the data.

Sample Characteristics:

NCVS target population includes people ages 12 and older living in households and group quarters within the U.S. and the District of Columbia. The sample of housing units is divided into 6 rotation groups, and each group is interviewed every 6 months for a period of 3½ years. In 2019, NCVS included 155,076 households and 249,008 people. The 2015 response rates for households and persons were 71 percent and 83 percent, respectively.

Alcohol Variables:

NCVS inquires if the victim noticed that the offender had been drinking or used drugs in combination with alcohol.

Other Variables:

NCVS includes demographic information on the victim and offender, characteristics of the crime, situational data, and information on responses by the victim about the incident and the criminal justice system. The recorded crimes (or attempted crimes) include rape and sexual attack, robbery, assault, pickpocketing, burglary, theft, motor vehicle theft, and vandalism.

National Emergency Medical Services Information System (NEMSIS)—2009–2020, Annually

Sponsoring Agency:

U.S. Department of Transportation National Highway Traffic Safety Administration, Office of Emergency Medical Services

Contact:

NEMSIS
P.O. Box 581289
Salt Lake City, UT 84158-1299
Telephone: 801-213-3930
nemsis@hsc.utah.edu

Availability:

Information on requesting the NEMSIS Public-Release Research Data set is at <https://nemsis.org/using-ems-data/request-research-data>. Obtain a thumb drive containing the data by submitting a request form online, via email nemsis@hsc.utah.edu, or post mail to NEMSIS Technical Assistance Center, University of Utah School of Medicine Department of Pediatrics, 295 Chipeta Way, P.O. Box 581289, Salt Lake City, Utah 84158-1220. The NEMSIS Data Cube, an online multidimensional database tool, is also available for collecting, storing, sharing, and analyzing NEMSIS data at <https://nemsis.org/view-reports/public-reports/ems-data-cube/>.

Overview:

NEMSIS is the repository for data on emergency medical services (EMS) data in the U.S. Standard data elements are collected by local EMS providers, then aggregated at the state level and submitted to the national database. Data from NEMSIS are used to evaluate care delivery, compare regional differences, inform EMS provider training, and generate research hypotheses.

Survey Design/Methodology:

Agencies submit data that characterize EMS ground activations (as opposed to patients receiving care) based on 9-1-1 requests for emergency care. Several states also submit data pertaining to interfacility/acute care transports and/or air medical transports. Process of care elements reported by EMS professionals are documented and used to characterize system responsiveness, medication/procedure use and effectiveness, protocol adherence, and other elements. Patient characteristics and detailed attributes of the underlying illness or injury are documented. The inclusion criteria that states use to characterize EMS activations aggregated in the state repositories vary. The data are limited by the absence of universal reporting.

Sample Characteristics:

The NEMSIS registry increased from 5,767,090 EMS activations submitted by 2,112 EMS

agencies serving 26 states in 2009 to 43,488,767 EMS activations submitted by 12,319 EMS agencies serving 50 states and territories in 2020. In 2012, 45 percent of all 9-1-1 EMS activations resulting in treatment and transportation involved patients who were ages 60 and older. The most frequent complaints reported to EMS by 9-1-1 dispatch centers were sick person (14 percent), breathing problem (10 percent), chest pain (8 percent), fall victim (8 percent), and traffic accident (5 percent). Provider's primary impression of the patient's condition most often cited was traumatic injury (13 percent), chest pain/discomfort (7 percent), abdominal pain/discomfort (7 percent), and respiratory distress (7 percent).

Alcohol Variables:

Alcohol use indicators in the EMS record include self-reported use, physical exam suspected use, and blood alcohol concentration.

Other Variables:

Substance use variables other than alcohol include opioid and other drug overdoses. Other items available for assessment include EMS agency characteristics, patient complaints reported to 9-1-1 call centers, delays experienced in the EMS response, identification of mass casualty events, and patient disposition decisions.

National Epidemiologic Survey on Alcohol and Related Conditions (NESARC)— Wave 1 (2001–2002), Wave 2 (2004–2005), Wave 3 (2012–2013)

Sponsoring Agency:

National Institute on Alcohol Abuse and Alcoholism (NIAAA), U.S. Department of Health and Human Services

Contact:

NESARC-III Data Access Committee
Laboratory of Epidemiology and Biometry
NIAAA, National Institutes of Health
6700B Rockledge Drive, Room 2127A
Bethesda, MD 20892
NIAAA-NESARC-III@mail.nih.gov

Aaron White, Ph.D.
Division of Epidemiology and Prevention
Research
NIAAA
Telephone: 301-451-5943
Whitea4@mail.nih.gov

Availability:

For confidentiality reasons, NESARC data have been designated as restricted access. Please contact Dr. White for access to Waves 1 and 2 data. Limited access to Wave 3 data sets is available through the Data Use Agreement; information on obtaining Wave 3 data is at <https://www.niaaa.nih.gov/research/nesarc-iii/nesarc-iii-data-access>.

Overview:

NESARC was designed to assess the prevalence of alcohol use disorders (AUD) and their associated disabilities in the general population. The survey is the largest ever comorbidity study of multiple mental health disorders among U.S. adults, including AUD, other substance use disorders, personality disorders, and anxiety and mood disorders. Wave 1 was fielded in 2001–2002. Wave 2 interviews were completed in 2004–2005 and used the same sample of respondents. Using a new sample of respondents, Wave 3 interviews were completed in 2012–2013.

Survey Design/Methodology:

NESARC is a nationwide household survey with a multistage stratified probability sample representative of civilian, noninstitutionalized adults residing in the U.S., including all 50 states and the District of Columbia. Military personnel living off base and residents in noninstitutionalized group quarters housing, such as boarding houses, shelters, and dormitories, were also included. One sample person age 18 or older was selected randomly from each household for a face-to-face interview. Data were collected using the computer-assisted personal interviewing method. Wave 3 also collected saliva samples from consenting respondents.

Sample Characteristics:

The final sample for Wave 1 included 43,093 respondents. Blacks, Hispanics, and young adults ages 18–24 were oversampled. The design and sampling strategy of the survey allow for population estimates at the national level. Wave 2

reinterviewed 34,653 of the 43,093 Wave 1 respondents. The final sample size for Wave 3 was 36,309 respondents.

Alcohol Variables:

Respondents were asked about their alcohol consumption behavior (e.g., drinking status, age of drinking onset, drinking and driving, and beverage-specific drinking amounts and patterns). Lifetime and past 12-month alcohol abuse (i.e., misuse) and dependence were measured by symptom questions according to the *DSM-IV* or *DSM-5* criteria for Waves 1–2 and Wave 3, respectively, using the NIAAA Alcohol Use Disorders and Associated Disabilities Interview Schedule–*DSM-IV/5* (AUDADIS-IV/5). Health and social consequences of drinking, alcohol treatment utilization, obstacles to obtaining treatment, and family history of alcoholism were also reported.

Other Variables:

Substance use variables other than alcohol include opioids, other illicit drugs, marijuana, prescription and non-prescription medicines, personal and family use of tobacco, dependence, and treatment utilization. Demographic variables include age, sex, race, Hispanic origin, childhood family structure, marital status, employment/school status, income, health insurance, selected medical conditions, and disability status. Mental health variables include lifetime and past 12-month DSM-(IV/5) diagnoses and treatment of major depression, dysthymia, mania and hypomania, panic disorder and agoraphobia, social and specific phobias,

generalized anxiety disorder, and pathologic gambling. Lifetime diagnoses were obtained for conduct disorder and 10 *DSM-IV* personality

disorders, including antisocial, avoidant, borderline, dependent, histrionic, narcissistic, obsessive-compulsive, paranoid, schizoid, and schizotypal personality disorders.

National Health and Nutrition Examination Survey I (NHANES I)—1971–1975

Sponsoring Agency:

National Center for Health Statistics (NCHS), Centers for Disease Control and Prevention

Contact:

Division of Health and Nutrition Examination Surveys

NCHS

3311 Toledo Road, Room 4551

Hyattsville, MD 20782-2064

Telephone: 800-232-4636

<https://wwwn.cdc.gov/nchs/nhanes/Default.aspx>

Availability:

Data files are available for download from <https://wwwn.cdc.gov/nchs/nhanes/nhanes1/default.aspx>. Due to confidentiality requirements, the NHANES I linked data files are only available for analysis through the NCHS Research Data Center (<http://www.cdc.gov/rdc/index.htm>).

Overview:

In 1970, the National Nutritional Surveillance System was combined with the National Health Examination Survey to form NHANES to initiate a series of surveys to collect information about health and diet of people in the U.S. Major goals of NHANES were to (1) estimate the number and percent of people in the U.S. population and designated subgroups with selected disease and risk factors; (2) monitor trends in the prevalence, awareness, treatment, and control of selected diseases; (3) monitor trends in risk behaviors and environmental exposures; (4) analyze risk factors for selected diseases; (5) study the relationship between diet, nutrition, and health; (6) explore emerging public health issues and new technologies; and (7) establish a national probability sample of genetic material for future genetic testing (NHANES III and beyond). NHANES I represents the first cycle of the NHANES studies.

Survey Design/Methodology:

NHANES I used a multistage, stratified probability sample of clusters of people ages 1 to 74, with oversampling of certain population subgroups—people living in poverty areas, women of childbearing age (ages 25–44), and elderly people (ages 65 and older). Data were weighted to represent the civilian, noninstitutionalized population of the U.S., excluding Alaska, Hawaii, and people residing on Tribal lands. During 1971–1979, extensive data were collected through interviews, physical examinations, a battery of clinical measures, and various laboratory tests. On the entire sample, these data include a general medical history; 24-hour dietary intake; food frequency interview; food program questionnaire;

general medical exam including dental, dermatological, and ophthalmological exams; anthropometric measures; and 24 hematological, blood chemistry, and urological lab determinations. Hand–wrist x-rays were performed on children ages 1–17, and additional clinical and laboratory tests were performed on a subset of sampled adults ages 25–74.

NCHS has conducted a linkage of NHANES I with records in the National Death Index (1971–2000), the Medicare Enrollment and Claims data (1991–2000), and the Social Security benefit history data (1962–2003). The linkage of the NHANES I survey participants with the other data provides opportunities to conduct studies designed to investigate the association of a variety of health factors with disability, chronic disease, health care utilization, morbidity, and mortality.

Sample Characteristics:

NHANES I sample included about 32,000 people ages 1–74. Among them, 14,407 people ages 25–74 were medically examined.

Alcohol Variables:

The NHANES I medical exam includes four alcohol questions:

During the last year, have you had at least one drink of beer, wine, or liquor?

How often do you drink?

Which do you most frequently drink (beer, wine, liquor)?

When you do drink (beer/wine/liquor), how much do you usually drink over 24 hours?

A 24-hour dietary recall questionnaire asks for the time and place of alcohol intake during a 24-hour period. Information on caloric value for each

ingested food substance is included. This permits analysis of food calories, alcohol calories, and percentage of alcohol in the respondent's diet.

Other Variables:

Demographic variables include age, sex, race, education, occupation, employment status,

marital status, income, language, and ancestry/national origin. Other variables include participation in public assistance programs, housing type and facilities, results of the medical history, 24-hour dietary intake, food frequency interview, and food program questionnaire, plus the general medical exams and laboratory tests.

National Health and Nutrition Examination Survey I Epidemiologic Follow-up Studies (NHEFS82, NHEFS86, NHEFS87, NHEFS92)

Sponsoring Agency:

National Center for Health Statistics (NCHS), Centers for Disease Control and Prevention

Contact:

Division of Health and Nutrition Examination Surveys
NCHS

3311 Toledo Road, Room 4551

Hyattsville, MD 20782-2064

Telephone: 800-232-4636

<https://wwwn.cdc.gov/nchs/nhanes/Default.aspx>

Availability:

Data files are available for download from <https://wwwn.cdc.gov/nchs/nhanes/nhefs/default.aspx>. Due to confidentiality requirements, the NHEFS linked mortality file is only accessible through the NCHS Research Data Center (RDC) at <http://www.cdc.gov/rdc/index.htm>.

Overview:

The NHEFS purpose was to investigate the relationships of clinical, nutritional, and behavioral factors assessed in NHANES I to subsequent morbidity and mortality. The three major objectives were to assess (1) morbidity and mortality associated with suspected risk factors, (2) changes in participant characteristics, and (3) natural history of chronic disease and functional impairments.

Survey Design/Methodology:

NHANES I respondents were traced and interviewed in 1982–1984 (NHEFS82), 1986 (NHEFS86), 1987 (NHEFS87), and 1992 (NHEFS92). Whereas NHANES I data include information gathered in physical exams, laboratory tests, and interviews, NHEFS was primarily a personal interview survey that relies on self-reporting of conditions. Interviews were conducted primarily by telephone. In addition, hospital and nursing home records were collected for any episode that occurred since the respondent's NHANES I examination, and death certificates were collected for those who had died. The sample is followed annually with the use of the National Death Index to obtain death certificates for respondents who have died between follow-up interviews. Health care facility records and death certificates were reviewed for the decedents. Pulse rate, weight, and blood pressure measurements were recorded for surviving participants.

Sample Characteristics:

Of the 14,407 respondents ages 25–74 examined in NHANES I, the numbers of living respondents

traced and completing interviews at follow-up surveys were NHEFS82 ($n = 10,523$), NHEFS86 ($n = 3,980$, ages 55–74 at baseline), NHEFS87 ($n = 11,018$), and NHEFS92 ($n = 10,079$).

Proxies were used for those who were incapacitated or had died.

Alcohol Variables:

NHEFS82 alcohol variables were derived from the following drinking questions:

Have you had at least 12 drinks of any kind of alcoholic beverage in any 1 year?

What is your main reason for not drinking?

Have you had at least one drink of beer, wine, or liquor during the past year?

What is your main reason for not drinking in the past year?

How old were you when you quit drinking?

On average, how often do you drink alcoholic beverages (i.e., beer, wine, or liquor)?

On days that you drink, about how many drinks do you usually have?

In how many of the past 12 months did you have at least 1 drink of any alcoholic beverage?

During the past 12 months, on about how many days did you have nine or more (five or more) drinks of any alcoholic beverage?

Do you now drink more, less, or about the same as you did a year ago?

Do you now consider yourself a light, moderate, or heavy drinker?

Which drinking category best describes your usual drinking pattern when you were 25, 35, 45, 55, 65, 75? For each age level, the categories range from

9+ drinks a day or 60+ per week to less than 1 drink a week.

Did you ever drink more than the amount you drank when you were (age of greatest drinking) for 3 months or longer? Which of the categories best describes your drinking during that period? About how old were you when you started drinking that amount? For about how long was this typical of your drinking?

NHEFS86, NHEFS87, and NHEFS92 contain subsets of these the variables derived from the above questions.

Other Variables:

Other NHEFS variables included demographics (age, sex, race, education, occupation, income, employment status, marital status); smoking and other tobacco use; medical history (medical conditions); functional limitation; exercise and weight; vision and hearing; pregnancy and menstrual history; urinary incontinence; changes in memory; nutrition (dietary recall and food frequency); physician's examination; living arrangements; household composition; community services; activity level; and utilization of hospitals, nursing homes, and other health care facilities.

National Health and Nutrition Examination Survey II (NHANES II)—1976–1980

Sponsoring Agency:

National Center for Health Statistics (NCHS), Centers for Disease Control and Prevention

Contact:

Division of Health and Nutrition Examination Surveys

NCHS

3311 Toledo Road, Room 4551

Hyattsville, MD 20782-2064

Telephone: 800-232-4636

<https://wwwn.cdc.gov/nchs/nhanes/Default.aspx>

Availability:

Data files are available for download from <https://wwwn.cdc.gov/nchs/nhanes/nhanes2/default.aspx>. Due to confidentiality requirements, the NHANES II linked data files only are available for analysis through the NCHS Research Data Center at <http://www.cdc.gov/rdc/index.htm>.

Overview:

NHANES II was designed to monitor the nutritional status and medical condition of the U.S. population. It consisted of eight elements, including questionnaires on household, medical histories for people ages 6 months to 11 years and ages 12–74, dietary intake, medication and vitamin usage, dietary supplement useage, and behavior. To establish a baseline for assessing changes over time, data collection for NHANES II was made comparable to NHANES I so the measurements for both surveys were taken in the same way and with the same age groups.

Survey Design/Methodology:

NHANES II employed a different sample design than that used with NHANES I. It used different definitions and stratification procedures to identify primary sampling units (PSUs). Three subgroups of the population were given special consideration in the nutritional assessment: preschool children (ages 6 months to 5 years), people ages 60 to 74, and people whose income was below the poverty level as defined by the 1970 U.S. Census. These procedures resulted in 64 PSU geographic locations.

NCHS conducted a linkage of NHANES II with records in the National Death Index (1976–2000) and the Medicare Utilization and Expenditure data (1962–2000). The linkage of the NHANES II survey participants with the other data provided opportunities to conduct studies designed to

investigate the association of a variety of health factors with disability, chronic disease, health care utilization, morbidity, and mortality.

Sample Characteristics:

NHANES II sampled 27,801 people ages 6 months to 74 years; 20,322 of them were given medical exams.

Alcohol Variables:

NHANES II included alcoholic beverage use in both the Dietary 24-Hour Recall and the Food Frequency questionnaire. Beer, wine, and liquor were included in the alcoholic beverages food group. The survey also had quantity–frequency questions covering a reporting window of 3 months. Drinking frequency response categories included never, less than once a week, and 1–6 times a week. Drinking quantity response categories included 1–24 times, 1–5 times, and 1–15 times per day.

Other Variables:

Demographic variables included age, sex, and race. Other variables include medical history, health history, dietary intake (24-hour recall and supplement), medications/vitamin usage, behavior questionnaire, control record, body measurements, audiometry, allergy testing, spirometry, liver function test, glucose challenge, speech pathology test, and physician's examination.

National Health and Nutrition Examination Survey III (NHANES III)—1988–1994**Sponsoring Agency:**

National Center for Health Statistics (NCHS), Centers for Disease Control and Prevention

Contact:

Division of Health and Nutrition Examination Surveys
NCHS
3311 Toledo Road
Hyattsville, MD 20782-2003
Telephone: 800-232-4636
<https://wwwn.cdc.gov/nchs/nhanes/Default.aspx>

Availability:

Data files are available for download from <https://wwwn.cdc.gov/nchs/nhanes/nhanes3/default.aspx>. Due to confidentiality requirements, the NHANES III linked data files only are available for analysis through the NCHS Research Data Center (<http://www.cdc.gov/rdc/>).

Overview:

NHANES III, the third cycle in the NHANES series, was conducted on a nationwide probability sample during 1988–1994. The survey was designed to collect information on the health and nutritional status of a national sample of the U.S. population through interviews and direct physical examinations.

Survey Design/Methodology:

NHANES III is the largest of the NHANES series so far. Black Americans and Mexican Americans were oversampled and comprised about 30 percent of the total sample. All selected participants were asked to complete an extensive interview and were examined in a mobile examination center. The survey period (1988–1994) consisted of two phases of equal length and sample size. Both Phase 1 and Phase 2 data collection involved random samples of the U.S. population living in households. NHANES III data were contained in five separate files (adult household, youth household, examination, laboratory, and dietary recall) that contain nearly all the data collected in the survey.

NCHS conducted a linkage of NHANES III with records in the National Death Index (1988–2000), the Medicare Enrollment and Claims data (1991–2000), and the Social Security benefit history data (1962–2003). The linkage of the NHANES III survey participants with the other data provides opportunities to conduct studies designed to investigate the association of a variety of health factors with disability, chronic disease, health care utilization, morbidity, and mortality.

Sample Characteristics:

NHANES III used a nationwide probability sample of 39,695 people ages 2 months and older, including large samples of both young and old respondents. About 12,000 of the sample were Black, 12,000 were Mexican Americans, and the remaining 16,000 were of all other race and ethnicity groups. In the 6-year sample, 33,994 sample participants were interviewed, and 30,818 participants were examined. NHANES III consists of 20,050 adult household data records, 29,314 lab data records, 13,994 youth household data records, and 31,311 examination data records.

Alcohol Variables:

Alcohol questions were asked of respondents ages 12 and older regarding alcohol use in the past 12 months, including number of drinking days, number of drinks per day on drinking days, number of days consumed 5+ and 9+ drinks, and ever consumed 5+ drinks almost every day (adults only). Frequency of drinking (beer, wine, hard liquor) in the past month were asked of participants ages 12 to 16 in the dietary food frequency section.

Other Variables:

Some of the 30 topics included high blood pressure, high blood cholesterol, obesity, passive smoking, lung disease, osteoporosis, HIV, hepatitis, helicobacter pylori, immunization status, diabetes, allergies, growth and development, blood lead, anemia, depression, food sufficiency, dietary intake, antioxidants, and nutritional blood measures.

National Health and Nutrition Examination Survey (Continuous NHANES)— 1999–2020, Biennially

Sponsoring Agency:

National Center for Health Statistics (NCHS), Centers for Disease Control and Prevention

Contact:

Division of Health and Nutrition Examination Surveys
NCHS

3311 Toledo Road, Room 4551

Hyattsville, MD 20782-2064

Telephone: 800-232-4636

<https://wwwn.cdc.gov/nchs/nhanes/continuousnhanes/default.aspx>

Availability:

Data files are available for download from

<https://wwwn.cdc.gov/nchs/nhanes/continuousnhanes/default.aspx>. Restricted data can be found through the NCHS Research Data Center at <https://www.cdc.gov/rdc/b1datatype/dt1222.htm>.

Overview:

The latest NHANES began in 1999 and became a continuous program focusing on various health and nutrition measures to meet emerging needs. The survey was designed to obtain nationally representative information on the health and nutritional status of the U.S. population through interviews and physical examinations.

Survey Design/Methodology:

The continuous NHANES examines a nationally representative sample of about 5,000 participants each year located in counties across the country, 15 of which are visited each year. All selected participants are asked to complete an extensive interview. More than 90 percent of those are given a physical examination in a mobile examination center (MEC). Audio computer-assisted personal self-interview and computer-assisted personal interview questionnaires are administered in the MEC. The data are contained in more than 50 separate files under the broad categories of demographic, examination, laboratory, and questionnaire.

Sample Characteristics:

NHANES samples include people in the civilian, noninstitutionalized population ages 2 months and older. Certain demographic subgroups, including Mexican American Hispanics, non-Hispanic Blacks, older adults, and low-income Whites/others are oversampled to enable accurate estimates for these groups. Starting in 2007, a new sampling methodology was implemented that oversampled all Hispanics, not just Mexican Americans. Starting in 2011, Non-Hispanic Asians were also oversampled. Beginning in 2007, the 12–15 and 16–19 age domains were combined,

and the 40–59 age minority domains were split into domains 40–49 and 50–59. Participation in laboratory tests depends on age and sex. Data are released in 2-year period cycles and are currently available for 1999–2000 through 2019–2020, with about 10,000 respondents per period.

Alcohol Variables:

Respondents ages 18 and older (20 and older before 2013–2014) were asked about their lifetime and past year alcohol use. Past-year questions included the number of drinking days, drinks per day on drinking days, and days consumed 5+ drinks. Beginning with the 2011–2012 survey, females were asked about consuming 4+ drinks instead of 5+. The amount of alcohol consumed in the past 24 hours and frequency of beer, wine, and liquor consumption in the past 30 days were also assessed. Respondents ages 12–17 (12–19 before 2013–2014) were asked about the number of days having 1+ drinks in life, as well as past-month alcohol use, including the number of days having 1+ drinks and 4/5+ drinks. These variables are not released to the public, but they can be analyzed through the NCHS Research Data Center.

Other Variables:

There are more than 50 topics investigated in the continuous NHANES. Substance use variables other than alcohol include tobacco, opioids, marijuana, and other drugs. Other topics include health insurance, immunization, physical activity, weight, dietary intake, reproductive history, sexual behavior, environmental exposures, physical fitness and functioning, mental health, cognitive functioning, hearing loss, vision, and select medical conditions.

National Health Interview Survey (NHIS)—General Description, 1997–2021, Annually**Sponsoring Agency:**

National Center for Health Statistics (NCHS), Centers for Disease Control and Prevention

Contact:

Division of Health Interview Statistics
NCHS

3311 Toledo Road, Room 2217

Hyattsville, MD 20782-2064

Telephone: 301-458-4901

nhis@cdc.gov

<http://www.cdc.gov/nchs/nhis.htm>

Availability:

Data files described at <https://www.cdc.gov/nchs/nhis/data-questionnaires-documentation.htm> are available for download from http://www.cdc.gov/nchs/nhis/quest_data_related_1997_forward.htm. Due to confidentiality requirements, the NHIS linked data files only are available for analysis through the NCHS Research Data Center (<http://www.cdc.gov/rdc/>).

Overview:

NHIS is a multipurpose health survey conducted continuously since 1957 by NCHS to obtain national information about the incidence and distribution of illness, its effects in terms of disability and chronic impairments, and the type of health services people receive. It is the principal source of health information on the U.S. civilian, noninstitutionalized, household population.

NHIS core questionnaire items are revised approximately every 10–15 years, with the last major revision occurring in 1997. In 1982–1996, NHIS consisted of two parts: (1) a core set of basic health and demographic items and (2) one or more supplemental sets of questions on current health topics. NCHS initiated a redesign of the NHIS questionnaire in 1997 to reduce the data collection burden and the interview length.

NHIS's redesign has three parts or modules: basic, periodic, and topical on prevention. The basic module functions as the new core questionnaire (consisting of four components: household, family, sample adult, and sample child). It will remain largely unchanged from year to year and will allow for trend analysis. For analytic purposes, data from more than 1 year can be pooled to increase the sample cell sizes.

Survey Design/Methodology:

NHIS is based on a stratified multistage sample design. Data are collected by the U.S. Census Bureau using computer-assisted interviews. For the family core component of the basic module, all adult members of the household ages 18 and older who are at home at the time of the interview are

invited to participate and to respond for themselves. For children and adults not at home during the interview, information is provided by a knowledgeable adult family member (ages 18 or older) residing in the household. From each family in the survey, one sample adult and one sample child (if any children younger than age 18 are present) are randomly selected. This adult responds to the questions in the sample adult questionnaire for him- or herself. Information for the sample child questionnaire is obtained from a knowledgeable adult in the household.

Changes in the state-level stratification increased the number of primary sampling locations from 198 to 358 in the 1995–2005 NHIS to enhance state estimation capabilities. Both Black and Hispanic populations are oversampled to allow for more precise estimation of health in these growing minority populations.

In 2006, a new sample design reduced the size of NHIS by about 13 percent relative to the previous sample design. Also starting in 2006, the new sample design included Asian respondents in the oversampling of minority populations. The sample adult selection process has been revised for the new sample design in 2006 so that Black, Hispanic, or Asian people ages 65 and older have an increased chance of being selected as the sample adult. In the 2016 design updates, sample addresses came from field listing on a limited basis and only in select areas. The NHIS's content and structure were updated with a new sampling scheme in 2019. One "sample adult" ages 18 or older and one "sample child" ages 17 or younger (if any children live in the household) are randomly

selected from each household. In March 2020, NHIS temporarily became a telephone survey due to the coronavirus pandemic.

Sample Characteristics:

Most NHIS families consist of a group of two or more related people living together in the same housing unit (household) in the sample. People living alone or unrelated people sharing the same household may also be considered as one family.

The sample sizes for 1997–2019 are as follows:

| Year | Families | Persons | Adults | Children |
|------|----------|---------|--------|----------|
| 1997 | 40,623 | 103,477 | 36,116 | 14,290 |
| 1998 | 38,773 | 98,785 | 32,440 | 13,645 |
| 1999 | 38,171 | 97,059 | 32,801 | 12,910 |
| 2000 | 39,264 | 100,618 | 32,374 | 13,376 |
| 2001 | 39,633 | 100,761 | 33,326 | 14,709 |
| 2002 | 36,831 | 93,386 | 31,044 | 12,524 |
| 2003 | 36,573 | 92,148 | 30,852 | 12,249 |
| 2004 | 37,466 | 94,460 | 31,326 | 12,424 |
| 2005 | 39,284 | 98,649 | 31,428 | 12,523 |
| 2006 | 29,868 | 75,716 | 24,275 | 9,837 |
| 2007 | 29,915 | 75,764 | 23,393 | 9,417 |
| 2008 | 29,421 | 74,236 | 21,781 | 8,815 |
| 2009 | 34,640 | 88,446 | 27,731 | 11,156 |
| 2010 | 35,177 | 89,976 | 27,157 | 11,277 |
| 2011 | 40,496 | 101,875 | 33,014 | 12,850 |
| 2012 | 43,345 | 108,131 | 34,525 | 13,275 |
| 2013 | 42,321 | 104,520 | 34,557 | 12,860 |
| 2014 | 45,597 | 112,053 | 36,697 | 13,380 |
| 2015 | 42,288 | 103,789 | 33,672 | 12,291 |
| 2016 | 40,875 | 97,169 | 33,028 | 11,107 |
| 2017 | 33,157 | 78,132 | 26,742 | 8,845 |
| 2018 | 30,309 | 72,831 | 25,417 | 8,269 |
| 2019 | n.a. | n.a. | 31,997 | 9,193 |
| 2020 | n.a. | n.a. | 31,568 | 5,790 |

Alcohol Variables:

Alcohol questions are in the NHIS core questionnaire and include the following: 12+ drinks in lifetime, 12+ drinks in the past year, frequency of drinking (number of days drank in the past year), average number of drinks on drinking days in the past year, and the number of days in the past year having had 5+ drinks for men or 4+ drinks for women. The 2019 redesigned NHIS annual core questionnaire did not ask alcohol questions. Alcohol questions continue to be asked every other year beginning in 2020.

Other Variables:

The survey includes tobacco and e-cigarette use variables since 2014, and other drug use variables through 2018. Other variables include many sociodemographic characteristics and variables related to limitation of activity, injuries, poisoning, health insurance, access to health care, health care utilization, health conditions, income and assets, immunizations, and testing for AIDS. The 2000 Cancer Control module covers Hispanic acculturation, diet and nutrition, physical activity, tobacco, cancer screening, genetic testing, and family history of cancer.

National Hospital Ambulatory Medical Care Survey (NHAMCS)—1992–2019, Annually

Sponsoring Agency:

National Center for Health Statistics (NCHS), Centers for Disease Control and Prevention, U.S. Department of Health and Human Services

Contact:

Ambulatory and Hospital Care Statistics Branch
NCHS
3311 Toledo Road
Hyattsville, MD 20782-2064
Telephone: 301-458-4600
<http://www.cdc.gov/nchs/about/major/ahcd/ahcd1.htm>

Availability:

Data files are available for download from https://www.cdc.gov/nchs/ahcd/data_sets_documentation_related.htm

Overview:

NHAMCS was initiated in late 1991 to fill the gap in coverage left by the National Ambulatory Medical Care Survey (NAMCS), which has collected data on ambulatory patient visits to physician offices since 1973. Part of the ambulatory component of the NCHS, NHAMCS is designed to gather, analyze, and disseminate information about the utilization and provision of ambulatory care services in hospital emergency and outpatient departments.

Survey Design/Methodology:

NHAMCS uses a national sample of visits to emergency and outpatient departments of noninstitutional general and short-stay hospitals, excluding federal hospitals, hospital units of institutions, and hospitals with less than six beds from the Verispan L.L.C., specifically Healthcare Market Index and Hospital Market Profiling Solution (formerly the SMG Hospital Market Database).

The survey uses a four-stage probability design with samples of geographically defined primary sampling units (PSUs), hospitals within PSUs, clinics within hospitals, and patient visits within clinics. The first-stage sample consisted of 112 PSUs that comprised a probability subsample of the PSUs used in the 1985–1994 NHIS. A fixed panel of 600 hospitals was selected for the NHAMCS sample. A total of 550 hospitals had an emergency department (ED) and/or an outpatient department (OPD), and 50 hospitals had neither an ED nor an OPD. From 2010–2012, NHAMCS also gathered data on ambulatory surgery procedures performed in freestanding ambulatory surgery centers. The

entire sample does not participate in a given year. Within each hospital, all outpatient clinics and emergency service areas (ESAs), or a sample of such units, are selected. Within ESAs or outpatient department clinics, patient visits are systematically selected over a randomly assigned 4-week reporting period. The actual visit sampling and data collection is primarily performed by hospital staff. The survey switched from paper to an automated laptop-assisted data collection method in 2012.

Sample Characteristics:

The NHAMCS basic sampling unit is the patient visit or encounter. In 2018, the sample included 20,291 patient record forms provided electronically by EDs.

Alcohol Variables:

ICD-9-CM and ICD-10-CM diagnosis codes are used to identify alcohol-related morbidity. The ED questionnaire asks whether the problem is alcohol related, and in 2014, alcohol abuse /misuse was added as a diagnosis. The outpatient questionnaire asks whether alcohol use counseling was ordered or provided.

Other Variables:

Drug misuse and related morbidity are identified by ICD-9-CM and ICD-10-CM diagnosis codes. In addition to demographics, patient information includes expected source of payment, major reason for visit, cause of injury, patient's complaint and symptoms, physician's diagnosis, and urgency of visit; services, procedures and medication ordered; referral status; and disposition of visit.

National Hospital Care Survey (NHCS)—2013–2016, Annually, and 2020–2021

Sponsoring Agency:

Centers for Disease Control and Prevention, National Center for Health Statistics (NCHS)

Contact:

NCHS 3311 Toledo Road

Hyattsville, MD 20782

Telephone: 301-458-4600

https://www.cdc.gov/nchs/nhcs/about_nhcs.htm

Availability:

NHCS data are available by proposal, including a fee submitted to the Research Data Center, with procedures described at <https://www.cdc.gov/rdc/index.htm>. COVID-19 hospital data are available without restrictions and may be downloaded or reviewed online at <https://data.cdc.gov/NCHS/COVID-19-Hospital-Data-from-the-National-Hospital-q3t8-zr7t>. Information regarding linked data, linkage methods, and related considerations is at <https://www.cdc.gov/nchs/data-linkage/nhcs-linkage.htm> or requested by email to datalinkage@cdc.gov.

Overview:

NHCS started data collection in 2011, integrating two long-standing NCHS surveys, the National Hospital Discharge Survey and National Hospital Ambulatory Medical Care Survey, and providing data on inpatient discharges and visits made to inpatient, emergency, outpatient, and ambulatory surgery center departments. Individual patients across settings and years can be linked for all data sources. Linkages of NHCS data have also been established with mortality information from the National Death Index, Medicare enrollment and claims data from the Centers for Medicare & Medicaid Services, and administrative housing assistance program data from the U.S.

Department of Housing and Urban Development. More specific linkages have been made with Drug-Involved Mortality Data and Enhanced Opioid-Identification Co-occurring Disorders among Opioid-involved Hospital Encounters. The linked data enable studies focused on the associations between a variety of health factors, health care utilization, and mortality. COVID-19 data are not nationally representative but may provide insight on the impact on various types of hospitals around the country.

Survey Design/Methodology:

The universe for NHCS consists of nonfederal and noninstitutional hospitals with six or more beds staffed for inpatient care in the 50 states and the District of Columbia. From 2013 through 2016, 581 hospitals were eligible to participate in the survey and/or complete the NHCS Annual Hospital Interview, which is offered to all hospitals in the sample regardless of participation status. Hospitals are randomly selected to provide

nationally representative data on hospital utilization. Participating hospitals are asked to submit all inpatient discharges, outpatient department encounters, and emergency department (ED) visits for up to a 12-month period. All NHCS data available in the Research Data Center are unweighted and are not nationally representative. Data from 2013 and 2014 are claims-based, and each record represents a unique encounter for a unique person, including the encounter's diagnoses, procedures, and reason for visit. Starting in 2015, electronic health record data are used, and one record is included for each diagnosis, procedure, or reason for visit.

Sample Characteristics:

From 2013 through 2016, there were approximately 125 million unique hospital encounters available, with 581 hospitals eligible to participate in the survey and/or complete the NHCS Annual Hospital Interview, which is offered to all hospitals in the sample regardless of participation status. The 2013 sample consisted of 1,474,478 inpatient discharges (97 hospitals), 3,784,397 ED encounters (82 hospitals), and 15,144,448 outpatient department encounters (87 hospitals). In 2014, 1,653,622 inpatient discharges (94 hospitals), 4,350,360 ED encounters (83 hospitals), and 19,005,777 outpatient department encounters (86 hospitals) were included. In 2015, 2,204,258 inpatient discharges (114 hospitals), 5,900,738 ED encounters (97 hospitals), and 26,455,149 outpatient department encounters (101 hospitals) were included. In 2016, 2,591,722 inpatient discharges (145 hospitals), 7,032,304 ED encounters (124 hospitals), and 35,692,420

outpatient department encounters (128 hospitals) were included.

The COVID-19 data, consisting of 1,192,211 ED encounters and 451,096 inpatient discharges in 2020 and 1,680,485 ED encounters and 536,517 inpatient discharges in 2021, are from 40 hospitals that submitted inpatient data and 40 hospitals that submitted ED data in 2020–2021. The NHCS data can show results by a combination of indicators related to COVID-19, such as length of inpatient stay, in-hospital mortality, comorbidities, and intubation or ventilator use. NHCS data allow for reporting on patient conditions and treatments within the hospital over time.

Alcohol Variables:

NHCS data include ICD-9-CM and ICD-10-CM diagnosis codes, which include alcohol-related and other substance use-related diagnoses.

Other Variables:

NHCS collects information from hospital settings for each encounter, including diagnoses, services, medications, laboratory tests and results (since 2015), discharge status, point of origin, and start and end date of the encounter. Variables for opioids and other recent substance use related to the patient's visit are included in the data set.

National Hospital Discharge Survey (NHDS)—1970–2010, Annually

Sponsoring Agency:

National Center for Health Statistics (NCHS), Centers for Disease Control and Prevention, U.S. Department of Health and Human Services

Contact:

Ambulatory and Hospital Care Statistics Branch
NCHS
3311 Toledo Road
Hyattsville, MD 20782-2064
Telephone: 301-458-4600
https://www.cdc.gov/nchs/nhds/about_nhds.htm

Availability:

Data and documentation files from 1970 and on are available for download from <https://www.icpsr.umich.edu/web/NACDA/series/43> or by ftp to ftp.cdc.gov/pub/Health_Statistics/NCHS/Data_sets/NHDS and [ftp.cdc.gov/pub/Health_Statistics/NCHS/Data set_Documentation/NHDS](ftp.cdc.gov/pub/Health_Statistics/NCHS/Data_set_Documentation/NHDS)

Overview:

NHDS was conducted continuously by NCHS from 1965 to 2010. NHDS annually abstracted both demographic and medical information from the face sheets of the medical records of inpatients selected from a national sample of hospitals. The survey was designed to provide national and regional estimates of hospital utilization by inpatients according to their demographic and medical characteristics as well as by characteristics of the hospitals, including their geographic location, bed size, and type of ownership. Beginning in 2011, previously collected NHDS inpatient data are integrated in the National Hospital Care Survey with emergency department, outpatient department, and ambulatory surgery center data collected by the [National Hospital Ambulatory Medical Care Survey](#).

Survey Design/Methodology:

NHDS covered discharges from community short-stay hospitals with an average patient length of stay of fewer than 30 days, general hospitals, or children's general hospitals, exclusive of federal or military hospitals, Veterans Administration hospitals, and hospitals with fewer than six beds located in the 50 states and the District of Columbia. In 1988, NHDS implemented a stratified, three-stage design in which units selected at the first stage of sampling consisted of either hospitals or geographic areas (i.e., 112 primary sampling units [PSUs] from the 1985–1994 National Health Interview Survey sample). Hospitals within PSUs were then selected at the second stage. Strata at this stage were defined by

geographic region, PSU size, abstracting service status, and hospital specialty-size groups. Within these strata, hospitals were selected with probabilities proportional to their annual number of discharges. At the final stage, a sample of discharges was selected by a systematic random sampling technique.

Sample Characteristics:

NHDS collected data from a sample of approximately 270,000 inpatient records acquired from a national sample of about 500 hospitals annually. Due to funding limitations, the sample of hospitals was reduced by half beginning in 2008. In 2010, the sample consisted of 239 hospitals. Of the 236 eligible hospitals, 203 hospitals responded to the survey. There were an estimated 35.1 million discharges of inpatients in 2010, based on 137,459 inpatient records (excluding newborn infants) from non-federal, short-stay hospitals.

Alcohol Variables:

NHDS diagnostic codes included those for alcohol-related morbidity (i.e., alcohol psychosis, alcohol dependence syndrome, cirrhosis of the liver, and nondependent abuse [i.e., misuse] of alcohol). ICD-9-CM codes were used.

Other Variables:

Each discharge record included the patient's demographic characteristics (sex, age, race, and marital status) and hospital characteristics (geographic region, ownership type and number of beds). Medical information included disease/injury diagnoses (up to seven per record), procedures performed (up to four per record), and discharge

status (dead or alive). Two additional items were included in the 2001 survey Medical Abstract form: type of admission and source of admission.

National Survey on Drug Use and Health (NSDUH)—2002–2020, Annually

Sponsoring Agency:

Substance Abuse and Mental Health Services Administration (SAMHSA), Centers for Disease Control and Prevention, U.S. Department of Health and Human Services

Contact:

Center for Behavioral Health Statistics and Quality
SAMHSA

5600 Fishers Lane, Parklawn Building
Rockville, MD 20857

Telephone: 240-276-1250 or Fax: 240-276-1260

CBHSQRequest@samhsa.hhs.gov

<https://www.samhsa.gov/data/data-we-collect/nsduh-national-survey-drug-use-and-health>

Availability:

Data files are available for download from <https://www.datafiles.samhsa.gov/data-sources> and from <https://www.icpsr.umich.edu/icpsrweb/ICPSR/series/64>. Online data analysis is also available on the website.

Overview:

The NSDUH series (formerly titled National Household Survey on Drug Abuse) is a major source of statistical information on the use of illicit drugs, alcohol, and tobacco and on mental health issues among members of the U.S. civilian, noninstitutional population ages 12 or older. The survey tracks trends in specific substance use and mental illness measures and assesses the consequences of these conditions by examining mental and/or substance use disorders and their treatments. The data are used to support prevention and treatment programs, monitor substance use trends, estimate the need for treatment, and inform public health policy.

Survey Design/Methodology:

NSDUH uses a multistage area probability sample of households and group quarters for each of the 50 states and the District of Columbia. The survey target population includes civilians living in households and certain noninstitutional group quarters (e.g., college dormitories, homeless shelters, and on military installations). It does not include military personnel on active duty and most transient populations, such as homeless people not residing in shelters. The survey uses computer-assisted personal interviews and audio computer-assisted self interviews

Sample Characteristics:

The sample sizes of public use data since 2002 are: 54,079 (2002), 55,230 (2003), 55,602 (2004), 55,905 (2005), 55,279 (2006), 55,435 (2007),

55,739 (2008), 55,772 (2009), 57,873 (2010), 58,397 (2011), 55,268 (2012), 55,160 (2013), 55,271 (2014), 57,146 (2015), 56,897 (2016), 56,276 (2017), 56,313 (2018), and 56,136 (2019). Sample weights are provided to permit national-level estimates.

Alcohol Variables:

NSDUH collects alcohol consumption information, including age at first use; most recent, lifetime, annual, and past-month use; number of days in the past month and past year on which respondents drank; number of drinks on days when respondents drank in the past month; and number of days respondents had 5 or more drinks in the past month. The threshold for binge alcohol use for females was lowered to 4 or more drinks on an occasion for the 2015 NSDUH. Thus, binge and heavy alcohol use among females are not comparable between 2015 and earlier years.

The survey questions allow for the operationalization of past-year *DSM-IV* alcohol abuse and dependence symptoms, criteria, and diagnosis, and the surveillance of other alcohol problems and treatment utilization for drinking. Starting in 2006, the survey incorporated a new consumption of alcohol module that collected additional information about respondents' last use of alcohol for those who indicated that they had consumed alcohol at least once in the past month. Other items in the consumption of alcohol module included the source of alcohol, location, and social context of the last drinking episode among past-month alcohol users ages 12 to 20; the number of drinks consumed on the

last drinking occasion; and the use of illicit drugs in combination with alcohol or within 2 hours of consuming alcohol on the last drinking occasion.

Other Variables:

Substance use variables other than alcohol include tobacco, opioids, marijuana, and other drugs. Related variables include beliefs concerning risk of various levels of use; symptoms of substance abuse (i.e., misuse) and dependence; overall health; and utilization of substance use treatment. Demographic variables include age, sex, race, education, occupation, and marital status. NSDUH also covers mental health, substance use treatment

history and perceived need for treatment, personal and family income sources and amounts, health care access and coverage, illegal activities and arrest records, problems resulting from the use of drugs, and needle sharing. NSDUH respondents are also asked for data concerning neighborhood environment; illegal activities; gang involvement; drug use by friends; social support; extracurricular activities; exposure to substance misuse prevention and education programs; and perceived adult attitudes toward drug use and activities such as school work, perceived risk of using drugs, perceived availability of drugs, driving behavior, and personal behavior.

National Mental Health Services Survey (N-MHSS)—2014–2020, Biennially; National Substance Use and Mental Health Services Survey (N-SUMHSS)—2021

Sponsoring Agency:

Center for Behavioral Health Statistics and Quality (CBHSQ) of the Substance Abuse and Mental Health Services Administration (SAMHSA), U.S. Department of Health and Human Services

Contact:

CBHSQ SAMHSA
5600 Fishers Lane, Parklawn Building
Rockville, MD 20857
Telephone: 240-276-1250 or Fax: 240-276-1260
CBHSQRequest@samhsa.hhs.gov
<https://www.dasis.samhsa.gov/dasis2/nmhss.htm>

Availability:

N-MHSS data files are available for download at <https://www.datafiles.samhsa.gov/data set/national-mental-health-services-survey-2020-n-mhss-2020-ds0001>.

Overview:

N-MHSS collects data annually on the services and characteristics of all known facilities in the 50 states, the District of Columbia, and the U.S. territories and jurisdictions that provide mental health treatment. It is the only source of national and state-level data on the mental health services delivery system reported by both publicly and privately operated specialty mental health care facilities. Every other year (since 2014), the survey also collects data on the number and demographics of people served in these facilities. Facilities are not eligible for inclusion if they only provide one or more of the following services: crisis intervention, psychosocial rehabilitation, cognitive rehabilitation, intake, referral, mental health evaluation, health promotion, psychoeducational, transportation, respite, consumer-run/peer support, housing, or legal advocacy. Residential facilities whose primary function is not to provide specialty mental health treatment services are also not eligible for inclusion. As of 2021, the National Substance Use and Mental Health Services Survey (N SUMHSS) (info.nsumhss.samhsa.gov) replaces the N-MHSS and the National Survey of Substance Abuse Treatment Services (N SSATS) by combining questions for substance use and mental health facilities into one survey.

Survey Design/Methodology:

The 2020 N-MHSS instrument is a 19-page document with 41 numbered questions. Topics include facility type, operation, and primary treatment focus; facility treatment

characteristics; pharmacotherapies; supportive services and practices; services provided to clients with co-occurring mental health and substance use disorders; dedicated or exclusively designed programs or groups offered; seclusion or restraint practices; crisis intervention team availability; and facility operating, management, and client characteristics. Data are collected three ways: a secure web-based questionnaire, a paper questionnaire sent by mail, and CATI.

Sample Characteristics:

The survey sample for the 2020 N-MHSS included 12,275 eligible respondent facilities. There were 77,622 and 43,744 clients who received mental health treatment services in inpatient and residential settings, respectively; and an estimated 3,593,843 clients received less-than-24-hour mental health treatment services in outpatient settings. Of 3,715,209 clients enrolled in mental health treatment services on April 30, 2020, an estimated 22 percent were diagnosed with co-occurring mental health and substance use disorders. Approximately 20 percent ($n = 2,478$) of facilities offered medication-assisted treatment for alcohol use disorder. The percentage of clients who had co-occurring mental health and substance use disorders varied by setting, from 20 percent in outpatient settings to 51 percent in facilities with inpatient, residential, and partial hospitalization/day treatment.

Alcohol Variables:

N-SUMHSS only includes questions regarding blood and urine testing of alcohol and

medication-assisted treatment for alcohol use disorder.

Other Variables:

Substance use variables other than alcohol include smoking/vaping/tobacco cessation counseling services, nicotine-replacement

therapy, and medication-assisted treatment for opioid use disorders.

Other variables include facility characteristics, mental health treatment modalities, client composition (e.g., sex, age, ethnicity, race, legal status), percent of clients with co-occurring substance use disorders, and smoking policies.

National Survey of Children's Exposure to Violence (NatSCEV)—1990–2008, 2011, and 2014

Sponsoring Agency:

United States Department of Justice. Office of Justice Programs. Office of Juvenile Justice and Delinquency Prevention (OJJDP)

Contact:

OJJDP
810 Seventh Street NW
Washington, DC 20531
Telephone: 202–307–5911

Availability:

Documentation may be downloaded, and restricted use data may be requested at <https://www.icpsr.umich.edu/web/NACJD/series/586>

Overview:

The NatSCEV was designed to obtain lifetime and one-year incidence estimates of a comprehensive range of childhood victimizations across sex, race, and developmental stage. It assessed the experiences of nationally representative samples of children ages 1 month to 17 years living in the contiguous U.S. (excluding New Hampshire) in three rounds of data collection: between January and July 2008 among children ages 1 month to 17 years for NatSCEV I; between March 2011 and January 2012 among children ages 1 month to 18 years for NatSCEV II; and between August 2013 and April 2014 among children younger than age 18 for NatSCEV III.

Survey Design/Methodology:

The samples consisted of two groups obtained through random digit dialing: a nationally representative sample of telephone numbers within the contiguous U.S. and an oversample of telephone exchanges with 70 percent or greater African American, Hispanic, or low-income households. Researchers conducted computer-assisted telephone interviews with youth ages 10 to 17 and with adult caregivers of children ages 9 and younger.

The survey used an enhanced version of the Juvenile Victimization Questionnaire, an inventory of childhood victimization that included these major areas of concern: conventional crime, child maltreatment, peer and sibling victimization, sexual assault, witnessing and indirect victimization (including exposure to community violence and family violence), school violence (including bullying) and threats, and internet victimization.

Follow-up questions for each victimization item included where the exposure to violence occurred,

whether injury resulted, how often the child was exposed to a specific type of violence, and the child's relationship to the perpetrator and (when the child witnessed violence) the victim. The questionnaire asked for household demographics and questions about the focal child's health.

Potential limitations of the survey were missing those children who were most vulnerable to being exposed to violence; underreporting or minimizing certain types of victimization by parents or caregivers who answer for younger children and might be unaware of the violence; missing or misclassification of some victimization episodes; and nonrecall by children of some exposure to violence or the timing of their exposure.

Sample Characteristics:

Sample sizes in the respective survey rounds were January–July 2008 ($n = 4,549$), March 2011–January 2012 ($n = 4,503$), and August 2013–April 2014 ($n = 4,000$).

Alcohol Variables:

Alcohol variables include lifetime alcohol drinking, current drinking, and alcohol use disorder in the family.

Other Variables:

The survey documents differences in exposure to violence across sex, race, socioeconomic status, family structure, region, urban/rural residence, and child's developmental stage; specifies how different forms of violent victimization "cluster" or co-occur; identifies individual, family, and community-level predictors of exposure to violence among children; examines associations between levels/types of exposure to violence and children's mental and emotional health; and assesses the

extent to which children disclose incidents of violence to various people and the nature and source of any assistance or treatment provided.

National Survey of Drinking and Driving Attitudes and Behavior—1991, 1993, 1995, 1997, 1999, 2001, 2004, and 2008

Sponsoring Agency:

National Highway and Traffic Safety Administration (NHTSA), U.S. Department of Transportation

Contact:

NHTSA Headquarters

1200 New Jersey Avenue SE.

West Building

Washington, DC 20590

Telephone: 888-327-4236

<https://www.nhtsa.gov/behavioral-research/behavioral-research-databases#national-telephone-surveys-20466>

Availability:

Data files for 2008 are available for download from <https://one.nhtsa.gov/Driving-Safety/Occupant-Protection/2008-National-Survey-of-Drinking-and-Driving-Attitudes-and-Behaviors>.

Overview:

NHTSA has conducted the National Survey of Drinking and Driving Attitudes and Behaviors periodically since 1991. The survey is designed to measure the scope of the drinking and driving problem and to guide efforts to reduce the problem's severity. The survey measures the status of attitudes, knowledge, and behavior of the general driving-age public about drinking and driving. Survey topics include frequency of drinking and driving, prevention and intervention, riding with impaired drivers, designated drivers, perceptions of penalties and enforcement, knowledge of BAC levels, and alcohol-impaired crashes.

Survey Design/Methodology:

The surveys were conducted in English or Spanish by telephone using a stratified Casady-Lepkowski Random Digit Dialing design and included only noninstitutionalized people in households with telephones. Both nondrivers and drivers were surveyed. In 1999, changes in sampling design were implemented to allow for state-level estimates.

Sample Characteristics:

The survey uses a nationally representative sample of the general driving-age public (ages 16 and older) selected by a multistage sampling procedure. In 1999, a requirement for a minimum of 100 completed interviews in each state and the District of Columbia was added. The final

numbers of record per year were: 1991 (2,406), 1993 (4,010), 1995 (4,008), 1997 (4,066), 1999 (5,127), 2001 (6,002), 2004 (6,049), and 2008 (6,999).

Alcohol Variables:

All versions of the drinking and driving survey include alcohol consumption items on frequency and usual quantity of alcohol consumption and beverage preferences. The 1993–1997 versions include graduated frequency items asking how often (1–2, 3–4, or 5+ times) drinks were consumed. The 1999 version has graduated frequency items asking how often (1+, 2+, 3+, 5+, and 8+) drinks were consumed. Beginning in 1993, all surveys also included the CAGE questionnaire that screens for alcohol problems. Drinking and driving variables included frequency of drinking and driving, frequency of driving while intoxicated, number of DWI convictions, frequency of riding with an impaired driver, support for taking action to reduce the problem, opinions about current enforcement and penalties, expectations of consequences, intervention behavior, and efforts by hosts to prevent guests from drinking and driving. Knowledge of BAC limits was added in 1995.

Other Variables:

Demographic variables include age, sex, race, income, education, employment status, and marital status.

National Survey of Family Growth (NSFG)—1973, 1976, 1982, 1988, 1995, 2002, 2006–2010, 2011–2013, 2013–2015, 2017–2019

Sponsoring Agency:

National Center for Health Statistics (NCHS), Centers for Disease Control and Prevention, U.S. Department of Health and Human Services

Contact:

NCHS
3311 Toledo Road
Hyattsville, MD 20782-2003
Telephone: 301-458-4222
nsfg@cdc.gov
https://www.cdc.gov/nchs/nsfg/about_nsfg.htm

Availability:

Public use data are available for download from https://www.cdc.gov/nchs/nsfg/nsfg_questionnaires.htm. Information on access to restricted data is at <https://www.cdc.gov/rdc/leftbrch/UseRestrictd.htm>.

Overview:

NSFG collects information on family life, marriage, divorce, pregnancy, infertility, use of contraception, and general and reproductive health. NSFG was first designed to be nationally representative of women ages 15–44 in the U.S. civilian, noninstitutionalized population. Later changes added an independent sample of men in 2002 and expanded the age range to ages 15–49 in 2015. NSFG can be used to provide reliable national trend data on substance use and pregnancy, such as smoking and alcohol use in pregnancy, and to estimate the number and characteristics of women in the U.S. at risk for an alcohol-exposed pregnancy.

Survey Design/Methodology:

The NSFG survey's first five cycles were conducted in 1973, 1976, 1982, 1988, and 1995, and were based on personal interviews conducted in the homes of a national sample of women ages 15–44 in the U.S. civilian noninstitutionalized population. Starting in 2006, NSFG shifted from a periodic survey to continuous interviewing. Interviews are done 48 weeks of every year for 4 years. Female interviewers are specifically trained to conduct NSFG using laptop computers (computer-assisted personal interviewing). The survey consists of three files: Female Respondent contains one record for each of the women interviewed; Female Pregnancy (interval) contains one record for each of the pregnancies reported by female respondents (pregnancy records are based on both completed

pregnancies [those that reached an outcome such as live birth, stillbirth, ectopic, miscarriage, or induced abortion] and current pregnancies [ongoing at time of interview]); and Male Respondent contains one record for each man interviewed. Individual-level variables that could not be included on the public-use NSFG files or could not be included in their original form are available to the research community through the NCHS Research Data Center.

Sample Characteristics:

The NSFG samples men and women ages 15–44 and oversamples Hispanics, Blacks, and teens. The 2013–2015 survey interviewed 10,205 participants.

Alcohol Variables:

Alcohol variables include the amount and frequency of alcohol consumption measured in the past 12 months. In the 2011–2013 survey, five new variables were added to assess alcohol consumption involving drinking over the past 30 days and binge drinking (defined as 4+ drinks on one occasion for females, and 5+ drinks for males). Females were asked if they believed alcohol consumption contributed to breast cancer. Participants were also asked if they were given alcohol during nonvoluntary sex.

Other Variables:

This survey can answer questions on factors affecting pregnancy, including sexual activity, contraceptive use, and infertility; factors affecting marriage, divorce, cohabitation, and family building; and attitudes about sex, childbearing, and marriage.

National Survey of Substance Abuse Treatment Services (N-SSATS)—2000, 2002–2019, Annually

Sponsoring Agency:

Substance Abuse and Mental Health Services Administration (SAMHSA), U.S. Department of Health and Human Services

Contact:

Center for Behavioral Health Statistics and Quality
SAMHSA

5600 Fishers Lane, Parklawn Building

Rockville, MD 20857

Telephone: 240-276-1250 or Fax: 240-276-1260

CBHSQRequest@samhsa.hhs.gov

<https://www.samhsa.gov/data/data-we-collect/n-ssats-national-survey-substance-abuse-treatment-services>

<https://www.samhsa.gov/data/data-we-collect/n-sumhss-national-substance-use-and-mental-health-services-survey>

Availability:

N-SSATS documentation and data files are available for download from

<https://www.dasis.samhsa.gov/dasis2/nssats.htm>, https://www.datafiles.samhsa.gov/data_set/national-survey-substance-abuse-treatment-services-2020-n-ssats-2020-ds0001,

and <http://www.icpsr.umich.edu/icpsrweb/ICPSR/series/58/studies?archive=ICPSR&sortBy=7>.

Overview:

N-SSATS is one of the three components of SAMHSA's Behavioral Health Services Information System (formerly the Drug and Alcohol Services Information System). The N-SSATS is a national survey designed to collect data on the location, characteristics, and use of substance use treatment facilities and services throughout the U.S., the District of Columbia, and other U.S. jurisdictions. The survey is used to assist state and local governments in determining the nature and extent of alcohol and drug treatment services provided at public, private, state-supported, and other treatment facilities. The survey also serves to help assess treatment resource needs; analyze and compare general treatment services on the national, regional, and state level; generate the *National Directory of Drug and Alcohol Abuse Treatment Programs*; and provide updated information for SAMHSA's Inventory of Behavioral Health Services (I-BHS—formerly the Inventory of Substance Abuse Treatment Services [I-SATS]) and the Behavioral Health Treatment Facility Locator database. The survey was formerly known as Uniform Facility Data Set (UFDS) (1995–1998) and the National Drug and Alcoholism Treatment Unit Survey (NDATUS) (1974–1994). As of 2021, the National Substance Use and Mental Health Services Survey (N-SUMHSS) (info.nsumhss.samhsa.gov)

replaces the N-SSATS and the National Mental Health Services Survey (N-MHSS) by combining questions for substance use and mental health facilities into one survey.

Survey Design/Methodology:

N-SSATS is a point-prevalence census and collects data from all active treatment facilities, including those on SAMHSA's I-BHS and those added by state substance use agencies. It uses three data collection modes: a secure web-based questionnaire, a paper questionnaire sent by mail, and a telephone interview and contains approximately 42 numbered questions.

Sample Characteristics:

A total of 15,961 providers responded to the survey in 2019. Facilities treating incarcerated persons only were identified and excluded in 2004.

Alcohol Variables:

Data are collected in three categories: drug, alcohol, and combined treatment services. This is a survey of facilities rather than patients, so it does not ask alcohol and/or drug questions per se. Data collected include unit orientation, types of alcohol/drug services offered, treatment modality and status, client characteristics, capacity and utilization on the point prevalence date, and payment source and fees charged. In the 2015 survey, 40 percent of the clients were

in treatment for both alcohol and drug misuse, and 15 percent were treated for alcohol only.

Other Variables:

Other variables include unit identification—location, type of environment, ownership, types

of programs and additional services provided, funding levels and sources, fees charged, hours of operation, and treatment capacity and utilization on the point prevalence date according to age, race/ethnicity, and sex by type of care by modality.

National Violent Death Reporting System (NVDRS)—2003–2019, Annually

Sponsoring Agency:

Centers for Disease Control and Prevention (CDC), U.S. Department of Health and Human Services

Contact:

National Center for Injury Prevention and Control

Division of Violence Prevention

CDC

4770 Buford Hwy, NE, MS F-63

Atlanta, GA 30341-3717

Telephone: 800-232-4636

dvpinquiries@cdc.gov

<https://www.cdc.gov/violenceprevention/datasources/nvdrs/index.html>

Availability:

NVDRS public use data are available for 2003–2005 at

<http://www.icpsr.umich.edu/icpsrweb/ICPSR/series/217>. Limited online analysis of data is available at:

<https://wisqars.cdc.gov:8443/nvdrs/nvdrsDisplay.jsp>. More detailed variables are included in restricted access data (RAD). Information on access requirements for RAD is available at

<http://www.cdc.gov/ViolencePrevention/NVDRS/RAD.html> and by email to nvdrs-rad@cdc.gov.

Overview:

NVDRS is a state-based surveillance system that links data on violent deaths collected from law enforcement, coroners and medical examiners, vital statistics, and crime laboratories. NVDRS's main objective is to assist in preventing violent deaths in the U.S. by providing systematically and routinely collected, accurate, timely, and comprehensive data for prevention program development. NVDRS's five main goals are to (1) collect and analyze timely, comprehensive data for monitoring the magnitude and characteristics of violent deaths at the national, state, and local levels; (2) ensure that violent death data are routinely and expeditiously disseminated to public health officials, law enforcement officials, policy makers, and the public; (3) track and facilitate the use of NVDRS data for researching, developing, implementing, and evaluating strategies, programs, and policies designed to prevent violent deaths and injuries at the national, state, and local levels; (4) build and strengthen partnerships with organizations and communities at the national, state, and local levels to ensure that data collected are used to prevent violent deaths and injuries; and (5) expand NVDRS in all 50 states, the District of Columbia, and U.S. territories.

Survey Design/Methodology:

NVDRS is a population-based, active surveillance system that provides a census of violent deaths that occur among both residents and nonresidents of funded U.S. states. NVDRS's coverage

increased from 6 participating states in 2002 to all 50 states and the District of Columbia and Puerto Rico as of 2018–2019. The CDC receives information about violent deaths from participating states' health departments. Cases consist of violent deaths these underlying causes (recorded in ICD codes): child maltreatment, suicide, homicide, undetermined intent, legal intervention, and unintentional firearm injury. Related fatal injuries involving multiple victims that occur within 24 hours of each other are linked in one incident.

Sample Characteristics:

The data include all violent deaths occurring in funded states so are not nationally representative. Data years and the number of states participating: 2002 (6 states), 2003 (13 states), 2004–2008 (16 states), 2009–2013 (18 states), 2014–2015 (32 states), 2016–2017 (40 states and the District of Columbia and Puerto Rico), and 2018–2019 (50 states and the District of Columbia and Puerto Rico).

Alcohol Variables:

The data set includes information on whether the victim's alcohol use was suspected, whether alcohol tests were conducted, the results of blood alcohol concentration tests, treatment utilization prior to death, and any possible alcohol problems preceding deaths for those who died of suicide or undetermined intents.

Other Variables:

NVDRS collects detailed information on victims and offenders, including demographics, toxicology testing and test results for substance use, manner of deaths, mechanism of injury, relationship of victim to offender, location of the incident (at home or work), date, type of incident, type of

weapon, and circumstances of the death. The circumstances of suicides and deaths of undetermined intent relate to mental health history and status, including whether the person disclosed intent to die by suicide, and other precipitating factors.

Panel Study of Income Dynamics (PSID)—1968–1996, Annually, 1997–2019, Biennially

Sponsoring Agency:

National Science Foundation

Contact:

Nancy A. Lutz
Directorate for Social, Behavioral, and Economic Sciences
National Science Foundation
2415 Eisenhower Ave, Alexandria VA 22314
Telephone: 703-292-7280
Nlutz@nsf.gov

Availability:

Public use data are available for download from <http://simba.isr.umich.edu/data/data.aspx> and <https://simba.isr.umich.edu/default.aspx>. Information on obtaining access to restricted data can be viewed at <http://simba.isr.umich.edu/restricted/ProcessReq.aspx>.

Overview:

PSID is designed to study the dynamics of income and poverty over the life course of families. PSID is the longest-running longitudinal household survey and collects information on employment, income, wealth, expenditures, health, marriage, childbearing, child development, philanthropy, education, and similar factors. PSID is the only data set to provide information on life course and multigenerational economic conditions, well-being, and health in a long-term panel representative of the full U.S. population.

Survey Design/Methodology:

From 1968 to 1972, more than 95 percent of the interviews were conducted face to face; since 1993, the survey has been administered using a computer-assisted telephone interview. PSID has released the main interview data in five data files: family, cross-year individual, birth history, marriage history, and parent identification. Multiple supplemental surveys were created, including the Child Development supplement and Transition into Adulthood supplement, Childhood Retrospective Circumstances study, Disability and Use of Time, and Wellbeing and Daily Life supplement.

Sample Characteristics:

PSID's sample include all people living in the selected families in 1968 plus anyone subsequently born to or adopted by a sample person. All sample persons are followed even after leaving to establish separate family units (FUs). This procedure replicates the

population's family-building activity and produces a dynamic sample of families each year. PSID families also include nonsample persons, most commonly people who marry sample persons after 1968. Information on nonsample persons is collected while they are living in the same FU as a sample person. Post-1968 immigrant families were added in 1997 to update the PSID by adding a representative sample of recent immigrants to the U.S., called the 1997 PSID Immigrant Refresher sample. More than 75,000 people have participated in PSID, and as many as six generations within sample families are represented. In 2019, 9,569 families were interviewed, totaling 26,084 participants.

Alcohol Variables:

Alcohol-related variables in the main survey include the frequency and amount of alcohol consumed in the last month and year, presence of an alcohol disorder, and whether or not the participant drank alcohol during pregnancy.

Other Variables:

Other variables include health status, onset of health conditions, health behaviors such as smoking and exercise, BMI, health insurance, and expenditures. Information about mental health was collected starting in 2001. Substance use variables other than alcohol include tobacco and e-cigarettes. A health history calendar was implemented starting in 2007 to collect information on early childhood health conditions, including age of onset and duration.

Population Assessment of Tobacco and Health (PATH)—2013–2021, Annually

Sponsoring Agency:

National Institute on Drug Abuse, National Institutes of Health (NIH), and the U.S. Food and Drug Administration (FDA), Center for Tobacco Products, U.S. Department of Health and Human Services

Contact:

PATHDataUserQuestions@Westat.com

Availability:

Public use files and documentation are available for download at <https://www.icpsr.umich.edu/web/NAHDAP/studies/36498>. Questionnaire and biospecimen-restricted use data are available by application through the ICPSR Access Restricted Data link at <https://www.icpsr.umich.edu/web/NAHDAP/studies/36231>. Researchers interested in accessing PATH study biospecimens for biospecimen research should refer to the PATH Study Biospecimen Access Program page.

Overview:

The PATH study is a nationally representative, longitudinal cohort study on tobacco use behaviors, including patterns of use, attitudes, beliefs, exposures, and health among the U.S. population. PATH is a collaboration between NIH, National Institute on Drug Abuse, and the the FDA Center for Tobacco Products designed to inform FDA's regulatory activities under the Family Smoking Prevention and Tobacco Control Act to reduce tobacco-related death and disease. The study sampled more than 150,000 mailing addresses across the U.S. to create a national sample of tobacco users and nonusers. Separate questionnaires are administered to adults, youth, and parents of youth. Adult and youth respondents are asked to complete an interview at each follow-up wave. Each adult interview respondent who completed the Wave 1 interview was asked to provide up to three biospecimens: urine, blood, and buccal cells. Youth in the study who turn 18 by the current wave of data collection are considered aged-up adults and are then invited to complete the Adult Interview. Additionally, shadow youth are considered aged-up youth upon turning 12 years old, when they are asked to complete an interview after parental consent. Data collection is planned through 2024.

Survey Design/Methodology:

The PATH population of interest is the 2013 civilian, noninstitutionalized U.S. population ages 12 and older. A four-stage stratified area probability sample design was used, with a two-phase design for sampling the adult cohort at the final stage. The Wave 1 weight for a PATH study respondent is a function of the inverse of the probability that a person is selected to be in the

PATH study sample; factors that adjust for nonresponse; and factors that calibrate estimates from the sample to quantities known from the 2010 U.S. Decennial Census and the 2013 American Community Survey. The study design oversamples tobacco users, young adults (ages 18–24) and African American adults. The study adapted many items from well-established existing national surveys, including the Tobacco Use Supplement to the Current Population Survey, NESARC, and NHANES. Questionnaire data are collected by bilingual field interviewers using computer-assisted self-interviewing and computer-assisted personal interviewing and various paper data collection forms.

Sample Characteristics:

A total of 45,971 respondents were interviewed in Wave 1 (September 2013–December 2014), including 32,320 adults administered the Adult Extended Interview and 13,651 youth administered the Youth Extended Interview. In addition, 13,588 respondents were administered the Parents Interview. Demographic features of respondents to the Adult Extended Interview were: 45.7 percent male, 54.3 percent female, 13.8 percent ages 18–24, 35.8 percent ages 25–44, 33.7 percent ages 45–64, 16.7 percent ages 65+, 14.1 percent Black alone or in combination, 77.4 percent white alone, 8.5 percent other race, 17.8 percent Hispanic, and 82.2 percent non-Hispanic. Demographic features of respondents to the Youth Extended Interview were: 51.2 percent male, 48.8 percent female, 34.3 percent ages 12–13, 66.7 percent ages 14–17, 28.6 percent Hispanic, 48.4 percent non-Hispanic white alone, and 23.0 percent non-Hispanic others. Wave 2 (October 2014–October 2015) consisted of 28,362 adults,

12,172 youth, and 12,129 parents of youth. Wave 3 (October 2015–October 2016) consisted of 28,148 adults, 11,814 youth, and 11,807 parents of youth.

Alcohol Variables:

Alcohol variables include lifetime use, current use, drinking frequency, and amount consumed per drinking occasion.

Other Variables:

Substance use variables other than alcohol include tobacco, marijuana, opioids, and other

drugs. The Adult Extended Questionnaire includes items regarding nicotine dependence, tobacco product packaging and health warnings, attitudes toward product regulation, media use, secondhand smoke exposure, peer and family influence, health effects, and exposure to industry advertising and promotion. The Youth Extended Questionnaire additionally includes items regarding tobacco product accessibility, psychosocial factors, and advertising exposure.

Research and Development Survey (RANDS)—RANDS 1 (2015), RANDS 2 (2016), RANDS 3 (2019)

Sponsoring Agency:

National Center for Health Statistics (NCHS), U.S. Department of Health and Human Services

Contact:

Research and Development Survey
NCHS

3311 Toledo Rd, Room 4635
Hyattsville, MD 20782-2064

RANDS@cdc.gov

<https://www.cdc.gov/nchs/rands>

https://www.cdc.gov/nchs/data/series/sr_01/sr01-64-508.pdf

Availability:

Public use data and documentation for RANDS 1, 2, and 3 may be downloaded from

<https://www.cdc.gov/nchs/rands/data.htm>. Each data file includes questionnaire data, demographic characteristics of the respondents, and sample weights. Data files are available in .csv and .txt formats. The codebook of variables, example SAS code to read the .txt file, survey questionnaire, and additional documentation are provided for each round.

Overview:

RANDS is an ongoing series of cross-sectional surveys sponsored by NCHS Division of Research and Methodology. These surveys are designed to explore the feasibility of using recruited commercial survey panels to collect information on national health outcomes and to augment NCHS's question evaluation and research program with quantitative methodologies for measuring error. Four rounds of RANDS surveys have been completed, with responses collected during fall of 2015, spring of 2016, spring of 2019, and summer of 2020, referred to as RANDS 1, RANDS 2, RANDS 3, and RANDS 4. In addition, NCHS adapted RANDS to collect timely, longitudinal information in three rounds on COVID-19 in the summer of 2020 and spring of 2021, referred to as RANDS during COVID-19. However, alcohol-related questions were not asked in RANDS 4 or RANDS during COVID-19 surveys.

Survey Design/Methodology:

The questionnaires were designed to be completed within 15 to 20 minutes and contain questions on health behaviors and conditions. RANDS 1, 2, and 3 collected responses using web administration, and RANDS 4 and RANDS during COVID-19 used web and phone administration. RANDS 1 and 2 (a focus on health conditions and behaviors) were conducted by Gallup. RANDS 3 and 4 (a focus on disability and opioids) and RANDS during COVID-19 surveys (a focus on COVID-19)

were conducted by the National Opinion Research Center at the University of Chicago. Each round was conducted as a probability sample using strata assigned by demographic factors such as race, ethnicity, age group, and education level. Poststratification weighting was used to maintain proportionality of demographic groups in the population. RANDS during COVID-19 rounds 1 and 2 surveys were also conducted as nonprobability samples. RANDS surveys contained existing questions from the National Health Interview Survey in addition to targeted, cognitive probe questions; several sets of experimental questions; and/or questions about COVID-19.

Sample Characteristics:

Each survey examined a sample of U.S. adults ages 18 and older with 2,304, 2,480, 2,646, and 3,442 respondents for RANDS 1–4, respectively.

Alcohol Variables:

RANDS 1 and 2 included lifetime and past-year drinking and high-intensity drinking questions. RANDS 3 includes a self-reported excess drinking question.

Other Variables:

The specific topics included on each round of RANDS varied, including questions on access to health care and utilization, chronic conditions, food security, general health, health insurance, physical activity, psychological distress, and disability. RANDS during COVID-19 was used to

publicly release a set of experimental estimates on select topics, including health status, chronic conditions, depression and anxiety, loss of work due to illness with COVID-19, health insurance and health care access, telemedicine access and

use, COVID-19–related health care and behaviors, and reduced access to health care. Substance use variables other than alcohol include smoking and opioid use.

Treatment Episode Data Set (TEDS)—1992–2019, Annually

Sponsoring Agency:

Substance Abuse and Mental Health Services Administration (SAMHSA), U.S. Department of Health and Human Services

Contact:

Center for Behavioral Health Statistics and Quality
SAMHSA

5600 Fishers Lane, Parklawn Building
Rockville, MD 20857

Telephone: 240-276-1250 or Fax: 240-276-1260

CBHSQRequest@samhsa.hhs.gov

<https://www.samhsa.gov/data/data-we-collect/teds-treatment-episode-data-set>

Availability:

Select documentation and data files are available for download from

<https://www.dasis.samhsa.gov/dasis2/teds.htm>. Data on admissions can be downloaded from

<http://www.icpsr.umich.edu/icpsrweb/ICPSR/series/56>, and data on discharges can be downloaded from

<https://www.icpsr.umich.edu/icpsrweb/ICPSR/series/238>.

Overview:

TEDS is one of the three components of SAMHSA's Behavioral Health Services Information System (formerly the Drug and Alcohol Services Information System). TEDS comprises two components: (1) the Admissions Data System (TEDS-A; data first reported in 1992), and (2) the Discharge Data System (TEDS-D; data first reported in 2000). TEDS collects information on the demographic, substance use, mental health, clinical, legal, and socioeconomic characteristics of people who are receiving publicly funded substance use and/or mental health services. The survey supports SAMHSA's initiative to build a national behavioral health data set accessible to the public; local, state, and federal policymakers; researchers; and others for examining comparisons and trends on the characteristics of people receiving substance use and/or mental health treatment services.

Survey Design/Methodology:

TEDS collects data on the number and characteristics of admissions to and discharges from state-administered public and private nonprofit substance use treatment programs in all 50 states, the District of Columbia, and Puerto Rico. State administrative data systems, claims, and encounter data are the primary data sources. TEDS provides a reporting framework for states to report treatment admissions and discharges of people receiving services. State representatives extract data from their states' system(s) to report to TEDS and, if needed,

convert state data elements to TEDS data definitions. Because significant differences exist among state data collection systems, state-to-state comparisons must be made with caution. TEDS includes a required minimum data set and an optional supplemental data set.

Sample Characteristics:

TEDS collects data from the states on admissions and discharges of people ages 12 and older. TEDS records represent admissions and discharges rather than individual patients, as a person may be admitted to or discharged from treatment more than once. TEDS does not include all admissions to and discharges from substance use treatment, but includes admissions to and discharges from facilities that are licensed or certified by a state substance use agency to provide substance use treatment. In general, facilities reporting to TEDS are those that receive state funds for the provision of treatment services, so TEDS does not represent the total national demand for substance use treatment. The TEDS system includes records for approximately 1.5 million substance use treatment admissions annually.

Alcohol Variables:

Patient alcohol use history, including frequency and age at first use, is collected along with clinical and treatment data such as service setting, number of prior treatments, referral data, diagnosis codes, and payment sources. In 2014, alcohol was the primary substance of use for 36 percent of all TEDS admissions.

Other Variables:

Other variables include patient demographics and history of heroin, marijuana, and other drug use.

Understanding America Study (UAS)—2014–2018, Ongoing

Sponsoring Agency:

U.S. Social Security Administration and National Institute on Aging, National Institutes of Health

Contact:

Understanding America Study
University of Southern California (USC)
PO Box 77902
Los Angeles, CA 90007
Telephone: 213-821-1819
uas-l@mymaillists.usc.edu
<https://uasdata.usc.edu/>

Availability:

A registration form for data access can be downloaded and submitted to uas-l@usc.edu at <https://uasdata.usc.edu>. The UAS Comprehensive File is a longitudinal aggregation of UAS core surveys. The longitudinal panel survey Drug Use Supplement (UAS-DUS) tracks monthly variation in drug use patterns and related characteristics. Further information about the UAS-DUS may be obtained by contacting the UAS or Dr. Junhan Cho at junhan.cho@usc.edu. A data visualization toolkit for interactively creating customizable and animated charts, tables, and maps can be accessed at <https://uasvis.usc.edu>.

Overview:

UAS is a probability-based nationally representative online panel of approximately 6,000 respondents ages 18 or older that began in 2014 administered by the Center for Economic and Social Research at USC. UAS's overall goal is to provide a detailed picture of people's daily lives both before and after retirement and other life events to enable the study of pathways to a wide range of outcomes. UAS includes more than 50 survey modules on topics such as retirement planning; financial and subjective well-being; health; social engagement; and various personality, cognitive, and other psychological constructs. It also includes modules that correspond topically with most of the modules that comprise the University of Michigan's Health and Retirement Study (HRS). The Understanding Coronavirus in America study was launched with a stand-alone survey on March 10, 2020, and included 29 waves of the tracking survey with a national long-form questionnaire and a Los Angeles County short-form questionnaire through July 20, 2021. Subsequent waves are conducted as stand-alone surveys. The USC Institute for Addiction Science commissioned the UAS-DUS resource, which integrates repeated measures of data collection that began in March 2020 during the Understanding Coronavirus in America Study and has subsequently continued with ongoing tracking surveys. Researchers may also commission access to the UAS panel, a large,

nationally representative sample for conducting customized survey-based and experimental research and analysis.

Survey Design/Methodology:

The target population of UAS surveys is typically noninstitutionalized U.S. residents ages 18 and older, although specific surveys may target other segments of the population. Panel members are selected and mailed surveys through address-based probability sampling sequential sample batching. Those who are female, middle-aged (ages 40 to 59), non-Hispanic, white, married, U.S. citizens, and U.S.-born are more heavily represented in the UAS panel than in the U.S. population. Each UAS survey is separately weighted using a two-step process to minimize the aggregate differences between the UAS panel and the U.S. population. Even after weighting, UAS panel members are slightly more likely than the U.S. population to have postgraduate education, and to be self-employed or unemployed, and differences exist in the distributions by citizenship, place of birth, and marital status. The UAS survey data set includes a set of standard variables, including individual and household identifying variables, demographic variables, and survey metadata. The core HRS instrument, which preceded the UAS, is administered every two years to enable direct comparability with the HRS. Important events in the lives of older UAS respondents are

monitored with brief, monthly assessments via the Internet, including anticipated events (e.g., retirement, job change, change in marital status) and unanticipated events (e.g., deaths, illnesses, job changes). Multiple domains of variables are tracked daily over an entire week, including daily pain, fatigue, physical functioning, stress, wellbeing, exercise, diet, social interaction, sleep, and cognitive function. Core COVID-19 survey modules were implemented as of March 2020, including questions in COVID-19-related domains: personal/family experiences, risk perceptions, protective behaviors, perceived effectiveness of protective behaviors, perceived safety/effectiveness of childhood vaccines, interest in receiving a COVID-19 vaccine, social interactions with family and friends, coping behaviors (substance use, binge drinking, and self-care), stress, anxiety depression, resilience, loneliness, and utilization of mental health care.

Sample Characteristics:

Demographic characteristics of the 5,310 UAS 2017 panel members were: female (56.1 percent); ages 18–39 (28.8 percent), 40–49 (19.2 percent), 50–59 (22.0 percent), 60 or older (30.0 percent); non-Hispanic white (78.1 percent), non-Hispanic Black (8.5 percent), Hispanic (7.1

percent), other race/ethnicity (6.3 percent); bachelor's degree or higher (35.7 percent); and income less than \$30,000 (27.1 percent).

Alcohol Variables:

Alcohol variables in the UAS Comprehensive file include ever drinking, frequency of drinking, and amount of drinking. Key measurements of the UAS-DUS include past-week use and frequency/intensity for alcohol, cannabis, cigarettes, e-cigarettes, and other drugs; depression symptoms; and anxiety symptoms.

Other Variables:

Domains included in the UAS Comprehensive file and core surveys are cognitive abilities, consumer behavior, crime, demographics, diet and lifestyle, education, employment and labor market, environment, family, financial literacy, health, health insurance, housing, income, non-cognitive abilities, politics, psychology, religion, retirement and pensions, risk preferences, savings, smoking, social attitudes and values, social networks, subjective expectations, subjective well-being, time preferences, and wealth. The UAS Comprehensive file also includes the harmonized version of the UAS-HRS.

Vital Statistics Mortality Data, Mortality Detail and Multiple Cause of Death —1968–2020, Annually

Sponsoring Agency:

National Center for Health Statistics (NCHS), U.S. Department of Health and Human Services

Contact:

Mortality Statistics Branch
Division of Vital Statistics
NCHS

3311 Toledo Road

Hyattsville, MD 20782

Telephone: 800-232-4636

<http://www.cdc.gov/nchs/deaths.htm>

Data Availability:

Multiple Cause of Death public use data files are available at

http://www.cdc.gov/nchs/data_access/Vitalstatsonline.htm#Mortality_Multiple. Online analysis of the Mortality Detail data for 1999–2020 is available at <http://wonder.cdc.gov/ucd-icd10.html>. Information on obtaining restricted data can be found at: https://www.cdc.gov/nchs/nvss/dvs_data_release.htm.

Overview:

The mortality data files contain information (e.g., demographic, cause of death, autopsy) from death certificates of all deaths occurring each year in the U.S. Using an ICD coding system, the Mortality Detail records only the underlying cause of death, and Multiple Cause of Death (MCD) records the underlying cause and up to 20 contributing causes. Mortality trend data are comparable with data from many countries as well as health-related data for small geographic areas in the U.S.

Survey Design/Methodology:

Data are collected from death certificates of 100 percent of reported deaths occurring in the U.S. each year (except for 1972, 1981, and 1982).

Sample Characteristics:

The total number of deaths varies from year to year. In 2019, about 2.9 million people died in the U.S. Deaths of nonresidents are excluded.

Alcohol Variables:

Some categories in the ICD are believed to be completely or nearly completely alcohol-related (i.e., alcohol psychosis, alcohol dependence

syndrome, nondependent alcohol abuse [i.e., misuse], and liver cirrhosis). These may be listed in records as underlying cause of death or as contributing cause of death (MCD only). In addition, research shows that other causes of death (e.g., suicide, homicide, motor vehicle crashes) often result from alcohol misuse in a different proportion of the cases. Using estimated fractions of alcohol's contribution to various causes of death, estimates can be derived on overall alcohol-related mortality.

Other Variables:

Substance use variables other than alcohol are based on ICD-9 and ICD-10 codes.

Demographic variables include sex, age, race and Hispanic origin, educational attainment, marital status, and residence. Death information includes direct underlying cause of death, contributing cause(s) of death (MCD only), autopsy findings, and date and place of death. Variables on county and actual date of death are restricted for reasons of confidentiality. Users must obtain special permission from NCHS to obtain these variables.

Section 2:

Special Population Data Sets

Adolescent Behaviors and Experiences Survey (ABES)—2021

Sponsoring Agency:

Division of Adolescent and School Health, National Center for HIV, Viral Hepatitis, STD, and TB Prevention, Centers for Disease Control and Prevention (CDC)

Contact:

CDC
1600 Clifton Rd.
Atlanta, GA 30329

Mike Underwood, Ph.D.
Chief, School-based Surveillance Branch
Division of Adolescent and School Health
Telephone: 404-718-1471
jmunderwoodncdc.gov

Availability:

ABES data and documentation can be downloaded at <https://www.cdc.gov/healthyyouth/data/abes/data.htm>. Selected results by demographic group are available at: <https://www.cdc.gov/healthyyouth/data/abes/tables/index.htm>

Overview:

CDC conducted ABES in spring 2021 to provide nationally representative data when many students were attending school virtually due to the Coronavirus Disease 2019 (COVID-19) pandemic. ABES was designed to: (1) assess the health impact of the COVID-19 pandemic, (2) determine the prevalence of health risk behaviors, and (3) examine the co-occurrence of health risk behaviors. The online questionnaire was designed to collect information on health-related experiences and behaviors among high school students, including: COVID-19–related experiences; emotional well-being; experiences related to perceived racism; behaviors that contribute to violence and unintentional injuries; sexual behaviors that contribute to unintended pregnancy and sexually transmitted infections, including HIV infection; alcohol and other drug use; tobacco use; unhealthy dietary behaviors; and inadequate physical activity.

Survey Design/Methodology:

ABES used a stratified, three-stage cluster probability-based sampling approach to obtain a nationally representative sample of students in grades 9–12 attending public and private schools. ABES used the same sampling methods as the national Youth Risk Behavior Survey (YRBS), except that it drew from a larger sample in anticipation of lower response rates. Combined data obtained from MDR (formerly Market Data Retrieval) and the National Center for Education Statistics were used to create the sampling frame. The sampling frame excluded alternative, special education, U.S. Department of Defense, Bureau of Indian Education, and vocational schools that serve students who are also enrolled in another public school. Home-

schooled students as well as students who dropped out of high school were not eligible for ABES if they were not enrolled in a school.

Because of the different instructional models used across the nation during the pandemic (i.e., in-person only, virtual only, and hybrid), ABES was designed as an online, self-administered, anonymous survey. The online administration allowed each school and teacher the flexibility to decide whether students should complete the survey during instructional time or on their own time. School-level data (e.g., instructional mode and school-provided equipment) were also collected. ABES data were weighted based on student sex and grade to account for school and student nonresponse and the oversampling of non-Hispanic Black students and Hispanic students. Students watched an online video with instructions for logging in and completing the questionnaire in English or Spanish.

The YRBS questionnaire, data collection procedures, and data processing were adapted to create ABES. The survey included 97 questions from the 2021 national YRBS questionnaire. Six of these questions were modified to allow students attending school only virtually to indicate a question asking about a behavior on school property did not apply. The questionnaire also included 12 questions not included on the YRBS questionnaire assessing COVID-19–related behaviors and experiences and one new question on perceived racism. ABES used photos and illustrations in connection with certain questions to improve students' recognition of particular tobacco products, drugs, and contraceptive methods. Each record contained information collected from the student's school, such as the school's

instructional model (e.g., in-person, virtual, hybrid).

Sample Characteristics:

Of the 339 sampled schools, 128 schools participated. Of the 16,037 sampled students, 7,998 submitted questionnaires; 7,705 questionnaires were usable after data editing. Among the high school students surveyed, 26.7 percent were in 9th grade, 25.5 percent in 10th grade, 24.3 percent in 11th grade, and 23.6 percent in 12th grade. In addition, 50.4 percent of students were female and 49.8 percent were non-Hispanic White, 25.4 percent were Hispanic, and 12.9 percent were non-Hispanic Black. The remaining students reported other races (all non-Hispanic): American Indian or Alaska Native (0.7 percent), Asian (4.9 percent), multiracial (5.8 percent), and Native Hawaiian or other Pacific Islander (0.5 percent). A majority of students self-identified as heterosexual (77.5 percent); 13.2 percent self-

identified as gay, lesbian, or bisexual; and 9.3 percent self-identified as other or questioning. Of the 128 participating schools, 75.0 percent reported using a hybrid instructional mode, 21.9 percent used online instruction only, and 3.1 percent used in-person instruction only. Approximately half the schools (50.8 percent) had reduced class sizes. Most schools (94.5 percent) provided laptops or Chromebooks to students, and 60.3 percent provided Wi-Fi hotspots.

Alcohol Variables:

Alcohol variables include current alcohol use, frequent current alcohol use, current binge drinking, frequent current binge drinking, largest number of drinks in a row, source of alcohol, and drank more alcohol during the COVID-19 pandemic.

Other Variables:

Students were asked if they used more drugs during the COVID-19 pandemic than prior to it.

American College Health Association-National College Health Assessment (ACHA-NCHA)—Fall 2000–Spring 2008; NCHA II—Fall 2008–Spring 2011; NCHA IIb—Fall 2011–Spring 2015; NCHA IIc—Fall 2015–Spring 2019; NCHA III—Fall 2019–Present

Sponsoring Agency:

American College Health Association (ACHA)

Contact:

Christine Kukich
ACHA
P.O. Box 28937
Baltimore, MD 21240-8937
Telephone: 410-859-1500
ckukich@acha.org
www.acha.org

Availability:

Biannually, ACHA compiles aggregate data from institutions using the ACHA-NCHA to provide a reference group for data comparison and publishes a Reference Group Report of these data. Proposals for access to portions of the Reference Group data are accepted; further information regarding data access is available by contacting Christine Kukich at ckukich@acha.org.

Overview:

ACHA originally developed ACHA-NCHA in 2000 to gain a comprehensive understanding of college students' health nationally and to offer individual higher education institutions the ability to identify the most pressing health issues facing their students, determine how to best allocate resources, gather formative data for campus health initiatives, inform policy decisions, track students' health over time, and benchmark against a national sample. NCHA examines issues relevant to the collegiate population, including alcohol and other drug use; sexual health; nutrition and exercise; mental health; and personal safety and violence. Participating schools can add their own supplemental questions.

Through fall 2019, 2 million students at 973 institutions have participated in ACHA-NCHA. More than half of these institutions have administered the ACHA-NCHA multiple times. Since its inception in 2000, a major overhaul of the survey occurred in 2008 and resulted in the ACHA-NCHA II. Minor revisions have been made since due to progressing terminology (e.g., identity categorizations for sexual orientation and gender identity) and emergent health issues (e.g., e-cigarettes). New in ACHA-NCHA III are questions related to parenting students, students in recovery, campus climate, food insecurity, homelessness, driving under the influence of cannabis, distracted driving, resiliency, flourishing, and firearms. NCHA III

also uses established, validated scales, including the Alcohol, Smoking, and Substance Involvement Screening Test, rather than a collection of individual items to assess several health issues, particularly within the mental health domain.

Survey Design/Methodology:

Sampling strategies used by schools included a mix of randomized classrooms, randomized mailings, or samples of all students; and randomized Web-based surveying. Domains of survey items in NCHA III were: overall health and community; weight, nutrition, and physical activity; sleep; safety; alcohol, tobacco, and other drugs; sexual health; well-being and mental health; health care utilization; health history; demographics; optional module – firearms.

Sample Characteristics:

Forty-two postsecondary institutions self-selected to participate in the fall 2021 ACHA-NCHA; all were four-year colleges or universities. Students completed 33,204 surveys on the 41 campuses that used random sampling techniques. Twenty-six campuses were public, and 15 were private. The numbers of students enrolled varied from fewer than 2,500 ($n = 6$) to 20,000 or more ($n = 16$). Eleven institutions were in the Northeast, seven in the Midwest, 17 in the South, and six in the West, providing a broad sampling of students across all U.S. regions. Five schools were in urban areas, and 13 were in urban settings with smaller

populations. Survey respondents reported 59 percent female, 41 percent male, and 3 percent transgender; median age 22.9; 92.3 percent full-time students, 53.8 percent taking in-person classes, 8.6 percent taking online classes, and 37.6 percent taking a mixture of in-person and online classes; 60 percent white, 6.4 percent Black, 1.7 percent American Indian/Native American, 18.7 percent Asian or Asian American, 4.1 percent multiracial, and 15.1 percent Hispanic.

Alcohol Variables:

Alcohol use variables include current use, number of drinks consumed during recent party

or social event, level of risk from alcohol use, driving after drinking, protective behaviors when drinking, and impact on academic performance.

Other Variables:

Substance use variables other than alcohol include tobacco, marijuana, opioids, and other drugs. Other variable domains include overall health and community; weight, nutrition, and physical activity; mental health; sleep; personal safety and violence; sexual health; well-being and mental health; health care utilization; health history; other academic impediments; demographics; and firearms (optional module).

Arrestee Drug Abuse Monitoring (ADAM I)—2000–2003; (ADAM II)—2007–2013, Annually

Sponsoring Agency:

Office of National Drug Control Policy (ONDCP), Executive Office of the President

Contact:

ONDCP
The White House
1600 Pennsylvania Ave., NW
Washington, DC 20500

Availability:

To access the restricted data, researchers must agree to the terms and conditions of a restricted data use agreement in accordance with existing inter-university consortium for political and social research servicing policies. To obtain information on how to access data, visit

<https://www.icpsr.umich.edu/web/ICPSR/series/110>.

Overview:

The original ADAM program was introduced in 2000 under the sponsorship of the National Institute of Justice. The original 35 counties in ADAM (2000–2003) were selected through a competitive grant process. In 2007, ONDCP reinstated the program as ADAM II and selected 10 counties from the original 35 based on geographic distribution to represent different regional drug use and adequacy of prior data. In 2012, ADAM II limited collection to 5 of the 10 counties. The retention of five sites was based on case production and response rates, cost efficiency, and geographic representation of drug use patterns. These sites represent the counties in which collection occurs through probability sampling of facilities and arrestees within those counties, but they cannot be used to generalize to national estimates. All instrumentation, sampling, and data collection protocols used in ADAM I have been replicated in ADAM II, permitting trend analysis from 2000 to 2013.

Survey Design/Methodology:

There are two levels of sampling: (1) sampling from the total number of facilities that book adult male arrestees in each county, and (2) sampling from the total number of adult male arrestees booked in a county. The sample is probability based and designed to represent adult male arrestees booked within 48 hours of arrest in the ADAM counties.

Data collection consists of (1) collection of booking information from official records, (2) a voluntary 20- to 25-minute face-to-face interview in the booking area of each facility, and (3) collection of a urine specimen. Respondents can complete the interview without providing a urine sample, but

they may not give a urine sample without completing the interview.

Sample Characteristics:

The 10 ADAM II sites for 2007–2011 were: Atlanta, GA (Fulton County); Charlotte, NC (Mecklenburg County); Chicago, IL (Cook County); Denver, CO (Denver County); Indianapolis, IN (Marion County); Minneapolis, MN (Hennepin County); New York, NY (Borough of Manhattan); Portland, OR (Multnomah County); Sacramento, CA (Sacramento County); and Washington, DC (District of Columbia). Beginning in 2012, five sites continued: Atlanta, Chicago, Denver, New York, and Sacramento. A total of 1,938 interviews were conducted and 1,736 urine specimens collected, weighted to represent more than 14,000 adult male arrestees.

Alcohol Variables:

Alcohol variables include the frequency of alcohol consumption and binge drinking (defined as 5 or more drinks on the same day), measured in the past 12 months and 30 days, and the purchasing of alcohol in the past 30 days. Additional alcohol variables include use of alcohol to relieve sadness, neglect of responsibility due to alcohol use, and objection of alcohol use from others. It was also reported if the first and second offenses involved possession of alcohol.

Other Variables:

Arrest information includes arrest date, time, location ZIP Code, three most serious charges and offense severity levels, and arrest and incarceration history. Sociodemographic variables include age, race/ethnicity, U.S. citizenship, education, employment, health insurance, marital status, and residency. Drug information includes urine drug test results; prior 3, 7, and 30 days

use; prior 12 months use by month; lifetime use; age at first use; method of obtaining drug; and drug and mental health treatment experience. Urine samples were tested in a central laboratory

for the presence of 10 drugs: marijuana, cocaine, opiates, amphetamine/methamphetamine, barbiturates, benzodiazepines, propoxyphene, PCP, methadone, and oxycodone.

CIRP Freshman Survey (TFS)—1966–2020, Annually

Sponsoring Agency:

Higher Education Research Institute

Contact:

Higher Education Research Institute
520 Portola Plaza, Math Sciences Room 4223
Box 951521
Los Angeles, CA 90095-1521
Telephone: 310-825-1925
heri@ucla.edu
<https://heri.ucla.edu/cirp-freshman-survey/>

Availability:

Select files are available for download from <https://heri.ucla.edu/data-archive>. Information on obtaining access to further data can be viewed at <https://heri.ucla.edu/data-access-for-researchers/>

Overview:

TFS is part of the Cooperative Institutional Research Program (CIRP), a national longitudinal study of the American higher education system. TFS is designed for incoming first-year students before they start classes at a college/university. TFS provides data on incoming college students' background characteristics, high school experiences, attitudes, behaviors, and expectations for college. Published annually in *The American Freshman*, the survey results provide a summary of the changing characteristics of entering students.

Survey Design/Methodology:

TFS is conducted before students start their college careers. Most campuses conduct the survey during orientation and use either paper or electronic formats. The best results occur when the survey is administered in a proctored setting. The paper survey can be administered in large-group settings during orientation, but it can also be provided to classrooms, residence halls, small groups, or through the mail. The web survey can be administered using an email

notification, managed either by the campus or Higher Education Research Institute.

Sample Characteristics:

TFS is designed for incoming freshmen before entering college. Reports provided to participating institutions can be broken down by sex, full and part-time student status, comparisons with other institutions, etc.

Alcohol Variables:

TFS asks students to indicate how often they drink beer, wine, or liquor: frequently, occasionally, or not at all.

Other Variables:

Other variables cover a wide range of student characteristics, including parental income and education, ethnicity, and other demographic items; financial aid; secondary school achievement and activities; educational and career plans; and values, attitudes, beliefs, and self-concept. The survey also addresses established behaviors in high school, academic preparedness, admissions decisions, expectations of college, interactions with peers and faculty, student values and goals, and concerns about financing college.

College Senior Survey (CSS)—1993–2020, Annually**Sponsoring Agency:**

Higher Education Research Institute

Contact:

Higher Education Research Institute © 2021
520 Portola Plaza, Math Sciences Room 4223
Box 951521
Los Angeles, CA 90095-1521
Telephone: 310-825-1925
heri@ucla.edu
<https://heri.ucla.edu/college-senior-survey/>

Availability:

Select files are available for download from <https://heri.ucla.edu/data-archive>. Information on obtaining access to further data can be viewed at <https://heri.ucla.edu/data-access-for-researchers/>

Overview:

CSS is part of the Cooperative Institutional Research Program (CIRP), a national longitudinal study of the American higher education system that is designed as an exit survey for graduating seniors. CSS focuses on a broad range of college outcomes and post-college goals, including academic achievement and engagement, student-faculty interaction, cognitive development, student goals and values, satisfaction with college, degree aspirations, and career and post-college plans.

Survey Design/Methodology:

CSS can be conducted using paper or electronic formats. The paper survey can be administered in the classroom, a group setting, or as a mail-out survey. The most successful way to administer CSS is in large group settings, such as a graduation rehearsal. The web survey can be administered through an email notification, managed either by the campus or Higher Education Research Institute. The web portal allows for email notification, reminder dates, and customizable pages, and colleges can upload additional questions.

Sample Characteristics:

CSS is designed to be answered by graduating college seniors. If seniors took the CIRP

Freshman Survey and were provided with a consistent ID number, their profiles can be matched and longitudinal reports can be produced, comparing freshman- to senior-year development. Reports can be broken down by sex, full- and part-time student status, comparisons with other institutions, etc.

Alcohol Variables:

CSS asks students how often they drink beer, wine, or liquor: frequently, occasionally, or not at all; and how many times in the past 2 weeks, if any, did they have 5 or more alcoholic drinks in a row: none, once, twice, 3–5 times, 6–9 times, or 10 or more times.

Other Variables:

Variables include past-year smoking; demographics (e.g., sex, race/ethnicity, religion, sexual orientation); involvement in school, research, extracurricular, and social activities; political views; and feelings regarding different aspects of college (e.g., class size, resources, advising). The survey also asks respondents to rate themselves on different characteristics (e.g., academic ability, physical health, self-confidence, tolerance) when compared with peers.

Department of Defense Health Related Behaviors Survey — 2015 and 2018. Department of Defense Health Behavior Survey—2008 and 2011. Worldwide Surveys of Alcohol and Nonmedical Drug Use Among Military Personnel—1980 and 1982. Worldwide Surveys of Substance Abuse and Health Behaviors Among Military Personnel—1982, 1985, 1988, 1992, 1995, 1998, 2002, and 2005.

Sponsoring Agency:

U.S. Department of Defense Military Health System

Contact:

Robert M. Bray, Ph.D.
RAND Corporation
Publications Orders, P.O. Box 2138
Santa Monica, CA 90407-2138
Telephone: 877-584-8642
order@rand.org

Availability:

Contact the RAND Corporation for information on data access.

Overview:

First administered in 1980, this survey was designed to measure prevalence of substance use and health behaviors among active-duty military personnel on U.S. military bases worldwide. Data can be combined to examine trends in substance misuse and negative effects of alcohol use from 1980 to 2008. Extensive changes to the methodology in 2011 preclude direct comparison to prior iterations of the survey. Changes include administration from group-administered paper-pencil to individual computer-based as well as sampling, weighting, data editing, and analysis. The 2005 survey introduced changes to questions related to illicit drug use by adding descriptions of drug use categories and revised the alcohol use items to be consistent with items from the Alcohol Use Disorders Identification Test.

Data are used to better understand the nature, causes, and consequences of substance misuse and health practices in the military and to help evaluate and guide related programs and policies. Comparisons between the military and civilian populations can be made using data from the National Survey on Drug Use and Health.

Survey Design/Methodology:

The survey collects data from all active duty military personnel worldwide in the five U.S. military services (Army, Navy, Air Force, and Marines; the Coast Guard was included beginning in 2008); a random sample is surveyed over a 6-

week period. Data are collected every 2–4 years, and more than 60 military installations worldwide are represented.

Sample Characteristics:

The final sample for 2018 consisted of 17,166 military personnel (3,646 Army, 3,675 Navy, 2,569 Marine, 5,579 Air Force, and 1,697 Coast Guard). Respondents were randomly selected to represent men and women in all pay grades.

Alcohol Variables:

The survey measures opioids, marijuana and other drugs, alcohol, and tobacco use in quantity and frequency during the past 30 days. Questions also cover negative physical, social, and work-related effects of alcohol and drug use as well as beliefs and attitudes about dangers related to use. Opinions about military alcohol and drug policies and programs are also reported.

Other Variables:

Other variables include positive health practices, knowledge/attitudes about AIDS, use of tobacco, exercise, diet, gambling, and injury prevention. The 1998 survey included questions for military women related to stress, coping styles, and special health issues. Questions were added in 2005 to better assess use of alternative medicine treatments, serious mental illness, and effects of deployment. The 2008 survey included new items geared toward characterizing deployment experiences and exposure to combat situations.

The Health and Retirement Study: A Longitudinal Study of Health, Retirement, and Aging (HRS)—1992–2020, Biennially

Sponsoring Agency:

National Institute on Aging, with supplemental support from the Social Security Administration

Contact:

HRS

Survey Research Center

426 Thompson Street

Ann Arbor, MI 48104

Telephone: 734-936-0314

hqsquestions@umich.edu

<https://hrs.isr.umich.edu/about>

<http://hrsonline.isr.umich.edu/help>

Availability:

Data files are available from <http://hrsonline.isr.umich.edu/index.php?p=avail>. Information on restricted data such as Medicare, earnings records, and geographic detail is available at

<http://hrsonline.isr.umich.edu/rda>.

Overview:

HRS is a nationally representative longitudinal panel study of older Americans' economic, physical, and mental health; marital and family status; and public and private support systems. It is designed to track age-related changes in health, economic status, and support that affect retirement, health insurance, saving, and well-being. A companion study, Assets and Health Dynamics Among the Oldest Old (AHEAD), is conducted in association with HRS to fill the gap of information on Americans older than age 70. HRS data can be linked with the Employer Pension Study (1993, 1999), the National Death Index, the Social Security Administration earnings and projected benefits data, W-2 self-employment data, and Medicare data.

Survey Design/Methodology:

The HRS core sample design is a multistage area probability sample of households. The baseline sample included in-home, face-to-face interviews in 1992 (1931–1941 birth cohort) and 1998 (1924–1930 and 1942–1947 birth cohorts). At six-year intervals, the six-year birth cohort that comprises ages 51–56 in that year is added to the sample. Follow-ups are conducted on these groups every two years, with proxy interviews after death. Blacks, Hispanics, and Florida residents are oversampled.

In 2006, HRS initiated an enhanced face-to-face interview. The enhanced interview includes the core interview plus a set of physical performance measures, collection of biomarkers, and a leave-behind questionnaire on

psychosocial topics. A random one-half of households were preselected for the enhanced face-to-face interview in 2006, with the other half of the sample selected for 2008. The design is repeated in each subsequent wave.

Sample Characteristics:

HRS currently surveys more than 22,000 Americans older than age 50 every two years. Each original sample is restricted to those living in households in the 48 conterminous states at the time of the baseline wave. Follow-up interviews are not restricted by geographic area. There are six distinct subsamples. The original HRS sample includes those born in 1931–1941. In 1998, the original AHEAD sample of those born before 1923 was merged with the HRS into a single interview schedule, and two groups were added: the War Baby sample, born in 1942–1947; and the Children of the Depression Age sample, born in 1924–1930. In 2004, the Early Baby Boomer cohort, born in 1948–1953, was added. The Middle Baby Boomer cohort, born in 1954–1959, was added in 2010.

Alcohol Variables:

Alcohol questions appear in the Health Status section and in the HRS Wave 1 Experimental Module on IADL Measures, Alcohol Use and History Module, and Substance Use Problems Module. Questions include lifetime use of alcohol, quantity and frequency of drinking in the past three months, opinions of what is considered a “drink,” and CAGE drinking problems (attempts to cut down, morning

drinking, and criticism and guilt about drinking).

Other Variables:

Substance use variables other than alcohol include tobacco, opioids, marijuana, and other drugs. Other variables include demographics; health status; health care utilization, cost, and

funding; cognitive conditions and status; attitudes, preferences, and expectations for the future; family structure and transfers; employment history; job demands and requirements; housing; income, assets, and net worth; disability; and health insurance and pension plans. Additional experimental modules are added each survey year.

Healthy Minds Study (HMS)—2015–2016, 2016–2017, 2017–2018, 2018–2019, 2019–2020, and 2020–2021

Sponsoring Agency:

Healthy Minds Network for Research on Adolescent and Young Adult Mental Health

Contact:

Liadan Solomon
 University of Michigan
 Research Data & Report Analyst
 Healthy Minds Network
 University of Michigan
 Telephone: 818-515-3514
liadan@umich.edu
hmn-datateam@umich.edu

Availability:

Download documentation and request the data sets for each year at:

<https://healthymindsnetwork.org/research/data-for-researchers/> or email healthyminds@umich.edu. An online interactive data interface for exploring the data is available at the same URL.

Overview:

The HMS survey assesses mental health, health service utilization, and related factors among college and university student populations. The survey gathers data on mental health status, access and barriers to services, utilization of services, social environment, academic environment, academic performance, and health behaviors (e.g., sleep and substance use). HMS has a special emphasis on understanding service utilization and help-seeking behavior, including factors such as stigma, knowledge, and the role of peers and other potential gatekeepers.

Survey Design/Methodology:

Approximately 200 participating campuses have implemented the web-based HMS. Each school provides a randomly selected sample of currently enrolled students ages 18 and older. Large schools typically provide a random sample of 4,000 students, and smaller schools typically provide a sample of all students. HMS has remained largely consistent, with minor changes each year to explore new topics or refine the measurement of existing topics. Three standard modules (Demographics, Mental Health Status, and Mental Health Service Utilization/Help-Seeking) form its core, and schools can elect to administer two or three elective modules from a menu of options (Knowledge and Attitudes about Mental Health and Mental Health Services, Upstander/Bystander Behaviors, Campus Culture and Climate, Resilience and Coping,

Sleep, Substance Use, Eating and Body Image, Sexual Assault, Overall Health, Academic Competition, Academic Persistence and Retention, Financial Stress, Diversity, Equity, and Inclusion, and Attitudes About Mobile Resources). Colleges and universities can also add up to 10 custom questions.

Sample Characteristics:

In 2020–2021, 55,553 students participated in the survey, consisting of 53 percent female, 45 percent male, and 2 percent other; 64 percent White, 13 percent Black, 13 percent Latino, 13 percent Asian, 2 percent Arab, 1 percent Pacific Islanders, and 2 percent Others; 39 percent living off-campus in nonuniversity housing; 29 percent living in campus residence, fraternity or sorority, or other university housing; and 28 percent in parent or guardian's home.

Alcohol Variables:

Alcohol use-related questions include current drinking, recent drinking, high intensity drinking, drinks per day, alcohol use disorders (AUD), AUDs Identification Test, perceptions of risk of high intensity drinking, other students' alcohol use, perceived problem with alcohol use on campus, alcohol-associated sexual assault, and upstander/bystander behavior.

Other Variables:

Substance use variables other than alcohol include smoking, vaping, marijuana, opioids, and other drugs. The 2020 survey included a COVID-19 module, with questions about COVID-19 diagnosis and symptoms and

COVID-19 among close family and friends;
social supports; precautionary and protective

measures; and related distresses and perceived
discrimination.

Hispanic Community Health Study/Study of Latinos (HCHS/SOL)—2008–2011, 2014–2017

Sponsoring Agency:

National Heart, Lung, and Blood Institute (NHLBI) and other institutes of the National Institutes of Health

Contact:

HCHS Administrator
Collaborative Studies Coordinating Center
Department of Biostatistics
Gillings School of Global Public Health
University of North Carolina at Chapel Hill
123 W. Franklin Street, Suite 450
Chapel Hill, NC 27516
HCHSAdministration@unc.edu
<https://sites.csc.unc.edu/hchs/>

Availability:

Use the form available at <https://biolincc.nhlbi.nih.gov/studies/hchssol/> to request a data package consisting of baseline and Visit 2 data from the main HCHS-SOL study as well as four ancillary studies: Sociocultural, Sueño, SOLNAS, and Youth.

Overview:

HCHS/SOL is a multicenter epidemiologic study in Hispanic/Latino populations that describes prevalence of major cardiovascular disease (CVD) risk factors and CVD among U.S. Hispanic/Latino people of different backgrounds, examines relationships of socioeconomic status and acculturation with CVD risk profiles and CVD, and assesses cross-sectional associations of CVD risk factors with CVD. HCHS/SOL addresses the prevalence of five major CVD risk factors (high serum cholesterol and blood pressure levels, obesity, hyperglycemia/diabetes, cigarette smoking), adverse CVD risk profiles, and coronary heart disease and stroke. Funding has been provided by the NHLBI and the National Institute on Deafness and Other Communication Disorders, National Institute of Dental and Craniofacial Research, National Institute of Diabetes and Digestive and Kidney Diseases, National Institute of Neurological Disorders and Stroke, National Center on Minority Health and Health Disparities, and NIH Office of Dietary Supplements.

Survey Design/Methodology:

HCHS/SOL examined self-identified Hispanic/Latino people ages 18 to 74 recruited from randomly selected households in four U.S. communities (the Bronx, Chicago, Miami, and San Diego). Households were selected using a stratified two-stage area probability sample design and were screened for eligibility.

During the baseline clinic visit in 2008–2011, study participants underwent an extensive clinical exam and assessments to determine baseline risk factors. Annual follow-up interviews are conducted to determine health outcomes of interest. During the 2014–2017 second clinic visit, participants were re-examined to collect data predictive of various health outcomes of interest. The average time between the baseline and second visit was six years. In addition, a comprehensive reproductive history of women of childbearing age was assessed. The third clinic visit phase began January 2020 and will conclude in early 2023.

Ancillary studies have been conducted involving the collection of additional data from HCHS/SOL participants or laboratory measurements from stored biospecimens collected by HCHS/SOL. The objective of HCHS/SOL Study of Latino Youth is to examine factors associated with childhood obesity and cardiometabolic risk among a diverse sample of Hispanic/Latino children living in one of the four HCHS/SOL U.S. communities.

The objective of HCHS/SOL: Nutrition and Physical Activity Assessment Study (SOLNAS) is to collect biological markers of dietary intake and physical activity for use in regression calibration models that will correct the bias of self-report and better estimate associations of dietary intake and physical activity with disease

outcomes. It is sponsored by the Albert Einstein College of Medicine Coordinating Center.

The objective of HCHS Sueño – Sleep Habits is to identify the specific psychosocial factors associated with altered sleep and the patterns of abnormal sleep associated with rigorously assessed cardiovascular outcomes within the Hispanic community.

The overall aims of the HCHS/SOL Sociocultural ancillary study are to examine the associations among socioeconomic status, sociocultural, and psychological risk and protective factors in relation to metabolic syndrome and CVD prevalence in a diverse Hispanic cohort using a unified theoretical framework.

Sample Characteristics:

Analyses involved 15,079 participants with complete data enrolled between March 2008 and June 2011. Participants included those of Cuban ($n = 2,201$), Dominican ($n = 1,400$), Mexican ($n = 6,232$), Puerto Rican ($n = 2,590$), Central American ($n = 1,634$), and South American ($n = 1,022$) backgrounds ages 18 to 74. In the study, 6,701 participants are in the 18-to-44 age group, 8,382 are in the 45-to-64 age group, and 1,332 are in the 65-to-74 age group. Sixty percent of the participants are female. Most participants completed some high school or graduated from high school, 16 percent graduated from college with a degree, and 23 percent attended at least

some college. Thirty-three percent have less than a high school education. Forty-two percent have a yearly household income of \$20,000 or less. Approximately three-fourths of the participants were not born in the United States. From December 2011 to December 2013, 1,466 participants ages 8–16 whose parents/legal guardians participated in HCHS/SOL were enrolled for SOL Youth. Of these, 1,129 children and 789 parent/primary caregiver participants from 775 households gave assent/consent for their data to be released for public use.

Alcohol Variables:

Alcohol variables include history of alcohol use, drinking patterns, current drinking, type of drinks, high intensity drinking, alcohol consumption, and alcohol use disorder.

Other Variables:

Substance use variables other than alcohol include tobacco and marijuana. Baseline exam components included a physical exam, blood samples, dental exam, hearing test, pulmonary function, physical activity assessment, and questionnaire data. Information obtained by questionnaires included demographic factors, socioeconomic status, acculturation (including years of residence in the U.S., generational status, and language preference), medications, sleep, respiratory/asthma, and medical history. Dietary intake was ascertained by two 24-hour dietary recalls administered six weeks apart.

Midlife in the United States Study 1, 2, 3 (MIDUS 1, 2, 3)—1995/1996, 2004–2005, and 2013–2014

Sponsoring Agency:

MacArthur Foundation Research Network on Successful Midlife Development; National Institute on Aging, National Institutes of Health, U.S. Department of Health and Human Services

Contact:

MIDUS – A National Study of Health and Well-Being
University of Wisconsin - Madison Institute on Aging
1300 University Avenue, 2245 MSC
Madison, Wisconsin 53706-1532
midus_help@aging.wisc.edu
<http://midus.wisc.edu/>

Availability:

Data files are available for download at: <https://www.icpsr.umich.edu/web/ICPSR/series/203/studies?archive=ICPSR&sortBy=7> and from the MIDUS Colectica Portal at <https://midus.colectica.org>, which has versatile tools that allow access to searchable variable-level metadata and descriptive statistics, the ability to find and merge variables across MIDUS projects, and the ability to create and download customized data set extracts with their corresponding codebooks.

Overview

MIDUS is a collaborative, multidisciplinary investigation of patterns, predictors, and consequences of midlife development. The primary objective is to identify the major biomedical, psychological, and social factors that permit some people to achieve good health, psychological well-being, and social responsibility during their adult years. MIDUS includes a core longitudinal survey administered in three waves in addition to several satellite studies that were conducted at various times. Cohort data are linked over time, and the study includes multiple cohorts and subsamples. MIDUS 1 was sponsored by the John D. and Catherine T. MacArthur Foundation. MIDUS 2, MIDUS Refresher, MIDUS 3, and the satellite surveys were supported by the National Institute on Aging.

Survey Design/Methodology

Data collection was performed through phone interview-administered surveys and mailed self-administered questionnaires at baseline and at various waves in addition to daily diary, cognitive, psychosocial, biomarker, and neuroscience assessments at various waves. Respondents were asked to provide extensive information on their physical and mental health throughout their adult lives and to assess the ways their lifestyles, including relationships and work-related demands, contributed to the conditions experienced.

MIDUS 1 was administered to a nationally representative sample of noninstitutionalized, English-speaking adults ages 25–74 obtained by random digit dialing (RDD) with oversampling of five U.S. metropolitan areas, older people, men, siblings of people from the RDD sample, and a national RDD sample of twin pairs. MIDUS 2 was conducted in 2004–2006 as a longitudinal follow-up. In 2011–2014, a national probability sample of adults ages 25–74, paralleling the five decadal age groups of the MIDUS 1 baseline survey, was recruited. The MIDUS Refresher was fielded in 2012 to replenish the original longitudinal sample with new participants. In 2013, a third wave (MIDUS 3) of survey data was collected on longitudinal participants then ages 40–94 years. In addition, random subsamples of respondents were recruited to participate in in-depth investigations of selected topics in key areas. These included: National Study of Daily Experiences (1996–1997); Biomarker Project (2004–2009, 2012–2016); Cognitive Project (2004–2006, 2011–2014, 2013–2017); Daily Stress Project (2004–2009, 2012–2014); Milwaukee African American Sample (2005–2006, 2012–2013, 2016–2017); Neuroscience Project (2004–2009, 2012–2016); Core Sample Mortality Data (2016); Boston Longitudinal Study of Cognition in Midlife (1995–2008); Psychological Experiences Follow-Up Study (1998); Survey of Minority Groups in Chicago

and New York City (1995–1996); and Midlife in Japan studies (2008).

Sample Characteristics:

MIDUS 1 main RDD participants ($n = 3,487$) were 49 percent male, ages 24–74 years, and 60 percent had more than 12 years of education. City oversamples ($n = 757$) were 57 percent males, ages 24–74 years, and 71 percent had more than 12 years of education. The sibling sample ($n = 950$) was 44 percent more than 12 years of education. The twin sample ($n = 1,914$) was 45 percent male, and 57 percent had more than 12 years of education. MIDUS 2 participants ($n = 4,963$) were 47 percent male, ages 32–84 years, and 67 percent had more than 12 years of education.

Alcohol Variables:

Alcohol variables in the main surveys and some satellite surveys included drinking history (i.e., age at first drink, age began drinking regularly), current drinking, drinking days, drinks per day, and high-intensity drinking.

Other Variables:

Substance use variables other than alcohol included tobacco, opioids, marijuana, and other legal and illegal drugs. These variables focused on history of use, regularity of use, attempts to quit, and effects of these substances on respondents' physical and mental well-being. Other variables included sense of control over health, awareness of changes in medical condition, regular exercise and a healthy diet, nontraditional remedies or therapies, attending support groups, overall well-being compared with peers, work histories, feelings of accomplishment, desire to learn, sense of control over their lives, interests, hopes for the future, quality of their childhood relationships with their parents and siblings, religion, rules/punishments, love/affection, and physical/verbal abuse. Demographic and background information included gender, age, education, marital status, income, and household composition. Geographic variables were not included.

Monitoring the Future (MTF): A Continuing Study of American Youth—1975–2021, Annually

Sponsoring Agency:

National Institute on Drug Abuse, U.S. Department of Health and Human Services, and Institute for Social Research, University of Michigan

Contact:

Survey Research Center
University of Michigan
P.O. Box 1248
Ann Arbor, MI 48106-1248
mtfinformation@umich.edu
<http://www.monitoringthefuture.org/>

Availability:

Data files are available for download from <https://www.icpsr.umich.edu/icpsrweb/ICPSR/series/35>. Online data analysis is also available at the website.

Overview:

MTF is designed to explore changes in important values, behaviors, and lifestyle orientations among contemporary American youth, with a particular emphasis on recent trends in licit and illicit drug use. Data have been collected each spring from high school seniors since the survey began in 1975; the study was expanded to include college students and young adults through follow-ups. Eighth- and 10th-grade students were added each year after 1990.

Survey Design/Methodology:

MTF employs a complex cohort sequential design appropriate for distinguishing and explaining period-related, age-related, and cohort-related changes. It can also be used to examine changes linked to different environments (e.g., high school, college, or employment) or role transitions (e.g., leaving the parental home, marriage, parenthood). The samples were drawn with a multistage random sampling procedure from public and private secondary schools throughout the conterminous U.S. The total 12th-grade sample was equally divided into six subsamples. Each subsample was administered a different form of the questionnaire to enable wide coverage of survey questions. However, about one-third of each questionnaire consists of the same core drug and demographic questions. The 8th- and 10th-grade surveys only used two different questionnaire forms in 1991–1996 (this expanded to four forms beginning in 1997). MTF's study design calls for biennial follow-ups—through age 32—of a subsample of the respondents in each participating senior class, beginning with the class of 1976.

Sample Characteristics:

Until 2020, approximately 50,000 8th-, 10th-, and 12th-grade students were surveyed each year from approximately 400 secondary schools. In 2020, the pandemic curtailed data collection; sample sizes were approximately 3,000, 5,000, and 4,000 for 8th, 10th, and 12th graders, respectively, from 112 secondary schools.

Alcohol Variables:

MTF includes lifetime, past year, and past 30-day alcohol use. Other alcohol questions include the grade during which a respondent first consumed alcohol; how many occasions respondents had been drunk or very high in their lifetime, past year, and past 30 days; the number of times they had either 4, 5, 10, or 15 or more drinks in a row in the last 2 weeks; and if they drank an alcoholic beverage containing caffeine or mixed with an energy drink in the last 12 months. Data are also collected on respondent attitudes and beliefs regarding alcohol and other drug use, perceived harm, perceived availability, and social disapproval. Eighth- and 10th-grade students are asked about the different locations in which they consume alcohol and if, in the last 2 weeks, they had been a passenger in a car where the driver had been drinking. Additional questions are asked for high school seniors, including simultaneous drug and alcohol use during the last 12 months. Variables for seniors also include the different occasions and reasons for alcohol use, treatment, and behavioral, health, and social problems resulting from alcohol use. Seniors are asked how many times they had driven a car, truck, or motorcycle after drinking in the last 2 weeks, and the number of times they received tickets or

warnings or had an accident while driving a car, truck, or motorcycle in the past 12 months after drinking alcoholic beverages.

Other Variables:

Substance use variables other than alcohol include marijuana, inhalants, hallucinogens, cocaine, heroin, other opiates, stimulants, sedatives, tranquilizers, cigarettes, cigars, hookah, smokeless tobacco, e-cigarettes, and

steroids. Sociodemographic data include sex, age, region, population density, and parental education, and other demographic and social network variables. A variety of other variables include information on attitudes toward religion, parental influences, changing roles for women, educational aspirations, self-esteem, social networks, exposure to sex and drug education, and violence and crime—both in and out of school.

National Child Abuse and Neglect Data System, Child File (NCANDS)—1995–2020, Annually

Sponsoring Agency:

Children's Bureau, U.S. Department of Health and Human Services

Contact:

National Data Archive on Child Abuse and Neglect

Beebe Hall—BCTR

Cornell University

Ithaca, NY 14853

Telephone: 607-255-7799

<http://www.ndacan.cornell.edu/about/about-contact-us.cfm>

Availability:

Because of the detailed nature of the information on child file records, these data are considered restricted. Researchers wanting to use the data must fulfill eligibility criteria, submit an application for approval to the archive, and enter into a legally binding data license that outlines the requirements for appropriate data use. Further information on access requirements is at

https://www.ndacan.cornell.edu/data_sets/data_sets-list-ncands-child-file.cfm.

Overview:

NCANDS is a federally sponsored, annual, national data collection effort created to track the volume and nature of child maltreatment reporting. The child file data set consists of child-specific data on all investigated maltreatment reports to state child protective service agencies. Beginning in 2000, the child file replaced the detailed case data component files, which included only data on substantiated or indicated maltreatment cases. State-level data are also available in a separate file.

Survey Design/Methodology:

States participate on a voluntary basis and submit their data after going through a process in which the state's administrative system is mapped to the NCANDS data structure. All reports reaching a disposition date (i.e., the report is completed) in a given year are mapped to the NCANDS data elements and included in the submission. Data are collected based on the federal fiscal year (FFY). The child file represents a census of all child protective services investigations or assessments conducted in the participating states. Individual child records are provided in a report.

Sample Characteristics:

Fifty states, the District of Columbia, and Puerto Rico submitted data to the NCANDS child file for FFY 2019. The resulting data set consists of 4,255,946 records.

Alcohol Variables:

The data set includes information on the child's compulsive use of or need for alcohol (including infants who are addicted at birth, are victims of Fetal Alcohol Syndrome, or may have other disabilities due to maternal use of alcohol during pregnancy), a caretaker's compulsive use of or need for alcohol that is not of a temporary nature, and whether substance use services were provided to the child and/or family.

Other Variables:

Other variables include the demographics of children and their perpetrators; types of maltreatment; investigation or assessment dispositions; risk factors for maltreatment, including compulsive drug use by the child and caretaker; and services provided as a result of the investigation or assessment, including mental health services.

National Longitudinal Study on Adolescent to Adult Health (Add Health)—Wave I (1994–1995), Wave II (1996), Wave III (2001–2002), Wave IV (2007–2008), Wave V (2016–2018), Parent Study (2015–2017)

Sponsoring Agency:

National Institute of Child Health and Human Development (NICHD) and 17 other federal agencies

Contact:

Add Health

Carolina Population Center

University of North Carolina at Chapel Hill, CB# 8120

206 West Franklin Street

Chapel Hill, NC 27516-2524

addhealth@unc.edu

<http://www.cpc.unc.edu/projects/addhealth/contact>

Availability:

Public-use data are available for download or online analysis at

<https://www.icpsr.umich.edu/web/DSDR/series/1006>; <https://dataverse.unc.edu/dataverse/addhealth> and

<https://addhealth.cpc.unc.edu/data/>. The more extensive restricted-use data are available by contractual agreement with the Carolina Population Center at <https://data.cpc.unc.edu/projects/2/view>.

Overview:

Add Health is a nationally representative study that explores the causes of health-related behaviors of adolescents in grades 7 through 12 and their outcomes in young adulthood. Add Health seeks to examine how social contexts (families, friends, peers, schools, neighborhoods, and communities) influence adolescent health and risk behaviors. In 2014, Add Health was renamed the National Longitudinal Study of Adolescent to Adult Health to reflect the study's ongoing nature that follows participants from early adolescence into adulthood. To date, data have been collected at five time points, Wave I (1994–1995), Wave II (1996), Wave III (2001–2002), Wave IV (2008–2009), and Wave V (2016–2018). The Add Health Parent Study gathered social, behavioral, and health survey data in 2015–2017 on a probability sample of the parents of the [Add Health](#) respondents originally interviewed in 1995.

Survey Design/Methodology:

The in-school phase (fall 1994) questionnaires were administered to students in 80 high schools and 52 associated middle schools identified through a stratified random sample of all high schools in the country. School administrators at each school completed a questionnaire on school characteristics and policies. In the in-home phases (Wave I, summer and fall 1995), interviews were conducted with a stratified sample of students enrolled in participating schools (core sample) and with selected oversampled students. A separate

interview was conducted with a parent of each adolescent in Wave I. Information about community and neighborhood characteristics were compiled independently from 1990 Census block group-level data and linked to the individual data. The in-home sample design includes a genetic sample of sibling pairs; a saturation sample of all adolescents attending selected high schools; a sample of students with disabilities; and an oversample of Chinese, Cuban, and Puerto Rican students and students from Black families with high levels of education. The Wave II in-home interview surveyed almost 15,000 of the same students one year after Wave I. The in-home Wave III sample consisted of Wave I respondents who could be located and re-interviewed 6 years later. Waves IV and V interviewed all eligible original Wave I in-home respondents available for in-home interviews.

Sample Characteristics:

Add Health includes 80 U.S. high schools and 52 middle schools with an unequal probability of selection. Systematic sampling methods and implicit stratification are incorporated into the study to ensure a sample representative of U.S. schools. At Wave I, 90,118 respondents participated in the in-school administration, and 20,745 respondents were interviewed in their homes. Of the respondents interviewed at home, 14,738, 15,197, 15,701, and 12,300 were reinterviewed at Waves II, III, IV, and V, respectively. At Wave V, the respondents were

between ages 32 and 42. Data for 2,013 Wave I parents surveyed in 2015–2017 are available in the Parent Study.

Alcohol Variables:

The in-home survey includes questions on alcohol consumption; binge drinking; perceived consequences of alcohol use; substance use in relation to driving, violence, and sexual behavior; and access to substances in the home. The parent survey includes questions on parent and child alcohol use.

Other Variables:

Substance use variables other than alcohol include tobacco, opioids, marijuana, and other drugs. The surveys asked questions about the student's daily activities, general health, self-esteem, personality, friends and peer networks, romantic relationships, pregnancy, contraception,

AIDS and STD risk perception, biological and resident parents, siblings, fighting and violence, delinquency, suicide, neighborhood, and religion.

The school administrator's survey asked questions concerning the school's characteristics, including type, specialization, class size, attendance level, teachers' sociodemographics and health-related behaviors, health education and services, SAT tests, and rules and discipline policies. A public use contextual database provides block group characteristics such as population, poverty, housing, education, labor force, and vital statistics. Wave V also included collection of blood pressure and pulse readings, anthropometric measures (height, weight, and waist circumference), venous blood for mRNA and DNAm, and other blood analytes from all consenting respondents.

National Longitudinal Survey of Youth (NLSY79)—1979–1994, 1995–2018, Biennially

Sponsoring Agency:

U.S. Department of Labor, National Opinion Research Center, and Center for Human Resource Research

Contact:

U.S. Bureau of Labor Statistics
National Longitudinal Survey Program
U.S. Bureau of Labor Statistics
2 Massachusetts Avenue, NE, Suite 4945
Washington, DC 20212-0001
Telephone: 202-691-7410
<https://www.bls.gov/nls/nlsy79.htm>

Availability:

Public use data are available for download from <https://www.nlsinfo.org/investigator/pages/search.jsp?s=NLSY79> and <https://www.bls.gov/nls/getting-started/accessing-data.htm>. Information about accessing the restricted data including geographical variables is available at <https://www.bls.gov/rda/home.htm> and <https://stats.bls.gov/nls/geocodeapp.htm>.

Overview:

NLSY79 was added in 1979 to the National Longitudinal Surveys (NLS) Series, sponsored by the U.S. Department of Labor, Bureau of Labor Statistics. NLSY79 is a national longitudinal survey to help evaluate the expanded employment and training programs for youth legislated by 1977 amendments to the Comprehensive Employment and Training Act. Since then, the NLSY has expanded to examine a variety of policy issues. The survey's aim is to obtain information on youth in the labor force and factors potentially affecting a young person's labor force attachment, including employment earnings, transition from school to work, training programs and training in the workplace, family/workplace relationships, geographic mobility, juvenile delinquency, and criminal behavior.

Survey Design/Methodology:

NLSY79 uses a multistage, stratified area probability sample designed to be representative of the noninstitutionalized civilian segment of American youth ages 14 to 22 when first interviewed in 1979. Supplemental samples oversampled civilian Hispanic, Black, and economically disadvantaged White youth. Another supplemental sample represented the military population ages 17 to 21. Annual personal interviews of the original respondents were conducted through 1994. Thereafter, interviews were biennial. The 1987 survey was conducted by phone. As of 2022, the survey consisted of 28 total rounds of data collection.

Sample Characteristics:

In 1979, NLSY79 sampled a total of 12,686 young people born between 1957 and 1964. This sample included 11,406 civilian and 1,280 military youth. Hispanic, economically disadvantaged, and youth in the military were oversampled. The respondents were ages 49 to 58 at the time of the 2014 interviews. After two subsamples were dropped, 9,964 respondents remain eligible for interview. Of 6,878 respondents in 2018, 51.9 percent were female, 49.0 percent were Non-Hispanic Non-Black, 31.5 percent were Black, and 19.5 percent were Hispanic or Latino.

Alcohol Variables:

Alcohol variables are included in the 1982–1985, 1988–1990, 1992, 1994, 2002, and 2006–2014 follow-up surveys. Questions provide information on drinking patterns, consumption of various alcoholic beverages, the impact of alcohol use on schoolwork and/or job behavior, frequency of going to bars, and trying to cut down on drinking. The 1988 survey included items about respondent relatives who have been alcoholics or problem drinkers.

Other Variables:

Substance use variables other than alcohol include tobacco, opioids, marijuana, and other drugs. Other variables in NLSY79 include demographics, marital history and fertility, education, labor force status, jobs and employer information, training, work experience and attitudes, military service, health limitations, and income and assets. Also included are questions on

job search methods, migration, educational and occupational aspirations and expectations, self-esteem, childcare, prenatal and postnatal health

behaviors, delinquency, time use, and AIDS knowledge.

National Longitudinal Survey of Youth 1979 (NLSY79) Child and Young Adult (NLSCYA) Child Sample—1986–2018, Biennially; Young Adult Sample—1994–2018, Biennially

Sponsoring Agency:

U.S. Bureau of Labor Statistics

Contact:

National Longitudinal Survey (NLS) Program
Postal Square Building, Suite 4945
2 Massachusetts Avenue NE
Washington, DC 20212-0001
Telephone: 202-691-7410 or 614-442-7366

usersvc@chrr.osu.edu

NLS_INFO@BLS.GOV

<https://www.bls.gov/NLS/>

Availability:

Instructions for accessing data are at <https://www.bls.gov/nls/getting-started/accessing-data.htm>. An online search and extraction site for reviewing NLS variables and creating data sets for each cohort is available at <https://www.nlsinfo.org/investigator>. Data files are available for download at https://www.bls.gov/nls/getting-started/nlscya_all_1979-2018.zip.

Overview:

In 1986, with funding from the National Institute of Child Health and Human Development and several private foundations, the NLS series was expanded to include surveys of a group of children born to women who participated in one of the national survey groups. The NLSY79 Child Survey, conducted jointly by the Ohio State University Center for Human Resource Research and National Opinion Research Center at the University of Chicago, is composed of all children born to mothers belonging to the NLSY79 cohort. Starting in 1994, NLSY79 children ages 15 and older formed the NLSY79 Young Adult sample, conducted biennially. As of 2016, children ages 12 to 14 have been included in the sample. Data are available from 1986 to 2018, representing 17 survey rounds for the child sample and 13 rounds for young adults.

Survey Design/Methodology:

The Child and Young Adult survey is modeled on the NLSY79 main youth questionnaire, tailored to this next generation, and designed for life course and cross-generational analyses. From 1986–1992, interviews with the NLSY79 children were conducted primarily in person using paper-and-pencil questionnaires. Beginning in 1994, the primary young adult and younger child instruments and assessments were administered using computer assisted personal interviewing. By 2002, all survey instruments were computerized. The size of the child and young adult samples increase over time, depending on the number of children born

to female NLSY79 respondents. School personnel completed a one-time school survey in 1995–1996 that contained information on each child's achievement, attendance, progress, activities, grades, and test scores. Geographic residence information is available in the main NLSY79 geocode data files for all children and young adults. Starting in 1986, the children of NLSY79 female respondents were assessed and interviewed every two years through 2014. In 2016, only the mother-reported assessments were completed as part of a mother supplement.

Sample Characteristics:

As of the 2018 interview round, the NLSY79 female respondents had attained the ages of 53 to 62 and had given birth to 11,545 children comprising 51 percent males and 49 percent females. Based on the mother's race/ethnicity, 53 percent were non-Black/non-Hispanic, 28 percent were Black, and 19 percent were Hispanic or Latino. It consisted of 5,255 children reported by 2,922 interviewed mothers in 1986; in 1994 (the first young adult survey year), 6,109 children younger than age 15 and 980 young adults were reported by 3,464 mothers. In 2016, interviewed NLSY79 mothers completed a limited number of questions about health and schooling for 236 children ages 18 or younger and for 4,965 children ages 12 years and older; nearly 80 percent of the young adult children fielded were interviewed as young adults. In addition, 3,011 mothers were interviewed, and 91 percent of their children have had at least one child interview, and 37

percent of ever-assessed children have data for all possible rounds between birth and age 14.

Alcohol Variables:

The survey includes questions regarding lifetime alcoholic beverage drinking; current drinking; peer pressure to drink; age at first use; the impact of alcohol use on school attendance and schoolwork; job behavior; peer drinking; problems with friends, family, neighbors, police and the law due to drinking; trying to cut down on drinking; and mother drinking during pregnancy.

Other Variables:

Substance use variables other than alcohol include tobacco, marijuana, heroin, and other drugs. Information collected includes education, training, employment, health, dating, fertility

and parenting, marriage and cohabitation, household composition, and social-psychological indicators. The young adult survey also includes questions on parent-child conflict, sexual activity, participation in delinquent or criminal activities, substance use, pro-social behavior, political attitudes, and their expectations for the future. Mother assessment measures include cognitive ability, temperament, motor and social development, behavior problems, and self-competence of the children as well as the quality of their home environment. Also collected from mothers are child demographic and family background characteristics, health information, and information on the child's home environment, including maternal emotional and verbal responsiveness and involvement with her child.

National Longitudinal Survey of Youth (NLSY97)—1997–2011, Annually, 2012–2020, Biennially

Sponsoring Agency:

U.S. Department of Labor, National Opinion Research Center, and Center for Human Resource Research

Contact:

National Longitudinal Survey Program
U.S. Bureau of Labor Statistics
2 Massachusetts Avenue, NE, Suite 4945
Washington, DC 20212-0001
Telephone: 202-691-7410
<https://www.bls.gov/nls/nlsy97.htm>

Availability:

Data available for download and online analysis is available at <https://www.nlsinfo.org/investigator/pages/search.jsp?s=NLSY97> and <https://www.bls.gov/nls/getting-started/accessing-data.htm>. Information about accessing the restricted data including geographical variables is available at <https://www.bls.gov/rda/home.htm> and <https://stats.bls.gov/nls/geocodeapp.htm>

Overview:

The National Longitudinal Survey of Youth 1997 (NLSY97) was added in 1997 to the National Longitudinal Surveys (NLS) Series, sponsored by the U.S. Department of Labor, Bureau of Labor Statistics. NLSY97 is designed to document the transition from school to work and into adulthood. It collects extensive information about labor market behavior and educational experiences among youth over time. Employment information focuses on two types of jobs: employee jobs, in which young people work for a particular employer, and freelance jobs, such as lawn mowing and babysitting. These distinctions enable researchers to study effects of very early employment among youth.

Survey Design/Methodology:

The first NLSY97 took place in 1997 and currently consists of 19 total rounds of data collection. In the 1997 round, both eligible youth and their parents underwent hour-long personal interviews, and an extensive two-part questionnaire was administered that listed and gathered demographic information on members of the youth's household and on his or her immediate family members living elsewhere. The young respondents are interviewed annually. Respondents self-administer potentially sensitive survey areas, such as those that address alcohol and drug use, sexual activity, and criminal behavior.

Sample Characteristics:

NLSY97 consists of a nationally representative sample of approximately 9,000 young people born between 1980 and 1984; at the time of first interview, respondents' ages ranged from 12 to 18. The respondents were ages 28 to 34 at their round 16 interviews (2013–2014). Two subsamples make up the NLSY97 cohort: a cross-sectional sample of 6,748 respondents designed to be representative of the initial survey respondents in 1980–1984, and a supplemental sample of 2,236 respondents that oversamples Hispanics and Blacks. A total of 6,734 respondents remain in the eligible 2018 samples.

Alcohol Variables:

Alcohol variables are included in the 1997–1998 and 1999–2020 follow-up surveys. Alcohol variables include lifetime and current drinking, age at first use, quantity, frequency, binge drinking (5+), drinking and driving, and drinking before or during work or school.

Other Variables:

Questionnaire subject areas include demographics, the relationships between youth and their parents, contact with absent parents, marital and fertility histories, dating, substance use, sexual activity, onset of puberty, training, participation in government assistance programs, expectations, time use, and criminal behavior. Substance use variables other than alcohol include tobacco, opioids, marijuana, and other drugs.

National Survey of Youth in Custody (NSYC)—2008–09; (NYSC-2)—2012; (NYSC-3)—2017–2018; National Survey of Youth in Custody, Alternate, Supplemental Survey on Drug and Alcohol Use (NYSC-A)—2008–2009, 2012, and 2018

Sponsoring Agency:

U.S. Department of Justice, Office of Justice Programs, Bureau of Justice Statistics (BJS)

Contact:

U.S. Department of Justice
Office of Justice Programs
BJS
810 7th St. NW, Washington, DC 20531
Telephone: 202-307-0765
askbjs@usdoj.gov

Availability:

Documentation and an application for use of the ICPSR Data Enclave can be download from pages associated with one of the data sets: e.g., <https://www.icpsr.umich.edu/web/NACJD/studies/33942>, <https://www.icpsr.umich.edu/web/NACJD/studies/37025>, or <https://www.icpsr.umich.edu/web/NACJD/studies/35039>

Overview:

NSYC is part of the BJS National Prison Rape Statistics Program to gather mandated data on the incidence of prevalence of sexual assault in juvenile facilities under the Prison Rape Elimination Act of 2003. The program collects administrative records of reported sexual violence as well as allegations of sexual victimization directly from victims through surveys of adult inmates in prisons and jails and surveys of youth held in juvenile correctional facilities in all 50 states and the District of Columbia. The universe for the survey was all adjudicated youth residing in facilities owned or operated by a state juvenile correctional authority and all state-adjudicated youth held under contract in locally or privately operated juvenile facilities. The surveys have been conducted by Westat, Inc., under a cooperative agreement with BJS.

Survey Design/Methodology:

Data are collected directly from youth in a private setting using audio computer-assisted self-interview technology. A multistage stratified sample design is used. NSYC and NYSC-2 each utilized two questionnaires based on the respondent's age. The Older Youth questionnaire was administered to respondents ages 15 and older, and the Younger Youth questionnaire was administered to those ages 14 and younger. NSYC is divided into five sections. Section A, Background, collected background information such as details of admission to facility and demographics, including education,

height, weight, race, ethnicity, sex, sexual orientation, and history of any forced sexual contact. Section B, Facility Perceptions and Victimization, included respondents' opinions of the facility and staff, any incidence of gang activity, and any injuries. Section C, Sexual Activity within Facility, captured the types and circumstances of sexual contact that occurred. Section D, Description of Events with Youth, and Section E, Description of Events with Staff Member, focused on when and where the contact occurred; the race and sex of the other youths or staff members; if threats or coercion were involved; and outcomes, including whether the sexual contact was reported.

NSYC-A is a supplement to NSYC. The survey was divided into six sections. Section A, Background, collected background information such as age, sex, education level, and whether respondent had stayed overnight in a facility or had forced sexual contact prior to current incarceration. Section B, Facility Perceptions and Victimization, is not included in this data set. Section C, Drug Use, included whether the respondent had ever used specific types of drugs, frequency of use in the past and immediately before being taken into custody, source of drugs, and symptoms of drug misuse and dependence. Section D, Alcohol Use, captured alcohol dependence and misuse symptoms. Section E, Treatment, focused on drug or alcohol treatment programs respondent had attended prior to being taken into custody.

Section F, Family and Peer Background, is not included in this data set.

The design of NSYC-3 was intended to obtain more details on specific incidents. New measures in NYSC-3 include items related to past physical and sexual abuse, youth's mental health and emotional problems, disabilities and other impairments, misconduct while in the facility, and placement in restricted housing. It is also intended to improve the measurement of the nature and circumstances surrounding staff sexual misconduct and boundary violations; collusion among inmates and staff surrounding victimization; impact on victims; and other factors related to facility climate, institutional culture, and correctional leadership.

Sample Characteristics:

Between June 2008 and April 2009, BJS conducted the first NSYC of 166 state-owned or operated facilities and 29 locally or privately operated facilities; a total of 10,263 youth participated. Of these, 1,065 received an alternative survey on drug and alcohol use and treatment, and 9,198 youth participated in the survey of sexual victimization. Data collection for NYSC-2 was conducted in 326 juvenile facilities between February and September 2012. A total of 8,845 youth completed the survey on sexual victimization after excluding 138 youth whose interviews were deleted due to extreme or inconsistent responses, and 996

completed the survey on drug and alcohol use and treatment. The NSYC-3 was conducted from March to December 2018 in 327 facilities that housed juveniles, including 217 state-owned or -operated facilities and 110 locally or privately operated facilities that held state-placed youth under contract. A total of 6,910 youth participated in the survey, with 6,211 completing the sexual-victimization survey and 699 completing an alternative survey on topics such as living conditions in the facility, mental health, drug and alcohol use, and education.

Alcohol Variables:

Alcohol variables include age at first drink; drinking history and frequency; alcohol use disorder symptoms; drinking-related problems; physical effects of drinking; whether offered drugs, alcohol, or cigarettes in the facility; and whether given alcohol or drugs by other youth or staff in the facility to facilitate sexual assault.

Other Variables:

Substance use variables other than alcohol include opioids, marijuana, other drugs, and related problems. Other variables include debriefing questions about respondents' experiences completing the survey, interviewer observations, variables to summarize victimization reports, weight and stratification data, and administrative data about the facilities.

NEXT Generation Longitudinal Study (NEXT Generation Health Study)—2009/2010–2016/2017**Sponsoring Agency:**

Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD), and other agencies of the U.S. Department of Health and Human Services

Contact:

NICHD

Social and Behavioral Sciences Branch
Division of Intramural Population Health
Research

Building 6710B, Room 3139D

Bethesda, MD 20892-7004

<https://www.nichd.nih.gov/about/org/diphr/hbb/research/Pages/next.aspx>

Denise L. Haynie, Ph.D., MPH

haynied@exchange.nih.gov

Telephone: 301-435-6933

Availability:

Data files are available for download at: <https://dash.nichd.nih.gov/study/15680>

Overview:

NEXT is a seven-year longitudinal assessment of a representative sample of U.S. adolescents and young adults starting at grade 10. The goals of NEXT include: identify the trajectory of adolescent health status and health behaviors from mid-adolescence through the post high school years; examine individual predictors of the onset of key adolescent risk behaviors and risk indicators during this period; identify genetic, personal, family, school, and social/environmental factors that promote or sustain positive health behaviors; identify transition points in health risk and risk behaviors and changes in family, school, and social/environmental precursors to these transitions; and examine the role of potential gene-environment interactions in the development of health status and health behaviors. Assessments were conducted annually for seven years beginning in the 2009–2010 school year.

Survey Design/Methodology:

In-school and online surveys collected self-reports of health status, health behaviors, and health attitudes as well as anthropometric data, genetic information, and neighborhood characteristics. The study incorporates a School Administrator Survey and other data files to obtain related information on school-level health programs and community-level contextual data. Additional data on BMI, pulse, and blood pressure were collected on a representative subsample of overweight and normal weight

adolescents (NEXT Plus sample); and driving performance was evaluated in 150 young adults.

Assessments were initiated with a nationally representative probability cohort of 10th-grade children in 2010, recruited using a multistage stratified design. Primary sampling units consisted of school districts or groups of school districts stratified across the nine U.S. Census divisions. African American youth were oversampled to provide a better population estimate and to provide an adequate sample to examine racial/ethnic differences in longitudinal predictors of health, health behaviors, and health behavior change. The cohort included an adequate sample of Hispanic youth to meet this criterion. In Wave 1, confidential self-report surveys were administered by trained research assistants in 10th-grade classrooms. Beginning in Wave 2, all longitudinal surveys were administered online or with hard copies when participants had limited online access.

Sample Characteristics:

Within the sampling framework, 137 schools were selected and formally recruited; 81 (59.1 percent) agreed to participate. Tenth-grade classes were randomly selected within each recruited school, and 3,796 students were recruited to participate; youth assent and parental consent were obtained from 2,874 (75.7 percent) (2,618 students in Wave 1 and an additional 256 students who were added in Wave 2). Of those who consented in Wave 1, 2,526 (96.5 percent) completed the Wave 1 survey. The Wave 1 sample included 57 percent

White, 20 percent African American, and 19 percent Hispanic respondents.

Alcohol Variables:

Alcohol variables included drinking prevalence, heavy episodic drinking, drinking venue, peer alcohol use, and drinking and driving.

Other Variables:

Substance use variables other than alcohol included tobacco, opioids, marijuana, and other

drugs. Baseline demographic variables included gender, race, family structure, and parental education as reported by the parent at Wave 1. Other data collected included parenting practices, depressive symptoms, residential status, school status, work status, social and leisure activities, physical activity, sleep behaviors, health status, and dating violence.

Nurses' Health Study (NHS)—1976–1988, Biennially; and 1989–2019, Annually

Sponsoring Agency:

Brigham and Women's Hospital, Harvard Medical School, and Harvard T.H. Chan School of Public Health

Contact:

NHS
Channing Division of Network Medicine
181 Longwood Avenue
Boston, MA 02115
Telephone: 617-525-2279 or Fax: 617-525-2008
nhsaccess@channing.harvard.edu

Availability:

Requests for collaboration and data access may be submitted by completing a form available at <https://nurseshealthstudy.org/researchers>. Questionnaires and further information about the data collected are available at the same web address.

Overview:

NHS was established in 1976 and has had continuous funding from the National Institutes of Health ever since. The study's original focus was on contraceptive methods, smoking, cancer, and heart disease, and expanded over time to include research on other lifestyle factors, behaviors, personal characteristics, and more than 30 diseases. Cohort members receive a follow-up questionnaire every two years with questions about diseases and health-related topics, including smoking, hormone use, and menopausal status. A food-frequency questionnaire for collecting dietary information was added in 1980 and continues to be mailed at four-year intervals. A quality-of-life supplement was first included in 1992 and is readministered at regular intervals. Supplemental questionnaires have been sent to selected participants to provide additional data and to better describe and define reported diseases. Blood, urine, toenail, and DNA samples have been collected at various times and stored and used to study the relationship between various biologic markers and disease risk.

Survey Design/Methodology:

The NHS cohort was established with a series of three mailings of the baseline questionnaire between June and December 1976 to married registered nurses (RNs) ages 30 to 55 who lived in the 11 most populous states (California, Connecticut, Florida, Maryland, Massachusetts, Michigan, New Jersey, New York, Ohio, Pennsylvania, and Texas). NHS 2 was established in 1989 with a younger cohort of

women ages 25 to 42 residing in 14 states (California, Connecticut, Indiana, Iowa, Kentucky, Massachusetts, Michigan, Missouri, New York, North Carolina, Ohio, Pennsylvania, South Carolina, and Texas) to study oral contraceptives, diet, and lifestyle risk factors. NHS 2 established a large biorepository beginning with blood samples collected in 1996, including samples collected during the luteal and follicular phases of the menstrual cycle. Further information was obtained on exposures in adolescence and early adult life, including physical activity, alcohol consumption, body fat profile, and diet. Participants were given the option to respond to web-based questionnaires beginning in 2001. To investigate factors that influence weight change, 27,805 children ages 9–14 of NHS 2 nurses were enrolled in their own follow-up study, the Growing Up Today study. This study had two enrollment waves: 1996 and 2004. NHS 3 was established in 2010 with participants including licensed vocation nurses, licensed practical nurses, and registered nurses ages 19–46 with the goal to include nurses from more diverse ethnic backgrounds, nurses in Canada, and male nurses. Beginning in 2015, recruitment was extended to male nurses. Every six months, NHS 3 participants complete a web-based questionnaire that collects data on current exposures and exposures during adolescence. In a substudy, participants report extensive information on exposures before, during, and immediately after pregnancy. Participant deaths and causes of death in all three cohorts are ascertained by

systematic searching of the National Death Index and state tumor registries.

Sample Characteristics:

There were 121,700 women enrolled in NHS 1. The cohort was 97 percent White, which reflected the racial composition of nurses at the time. NHS 2 enrolled 116,430 women. NHS 3 has enrolled more than 40,000 female and male nurses ages 19 to 49 residing throughout the U.S. and Canada, with 14 percent self-identifying as members of a racial or ethnic minority.

Alcohol Variables:

NHS questionnaires since 1988 include items for reporting amounts and frequencies of alcohol consumption by type of alcoholic beverage (i.e., beer, light beer, white wine, red wine, and liquor).

Other Variables:

Other variables included behavioral and lifestyle risk factors: smoking, leisure time, physical activity, sedentary time, sleep patterns, weight, height, waist and hip measurements, diet including during adolescence with a semiquantitative food frequency questionnaire, employment status, reproductive history and menopause, prescription and over-the-counter medications, cancer and other screening tests, living arrangement, neighborhood characteristics, environmental exposures, mental health, social networks, optimism scale, caregiving and caregiving stress, quality of life, activities of daily living, and personal and family medical history.

The Pregnancy Risk Assessment Monitoring System (PRAMS)—Phase I (1988–1989), Phase II (1990–1995), Phase III (1996–1999), Phase IV (2000–2003), Phase V (2004–2008), Phase VI (2009–2011), Phase VII (2012–2015), and Phase VIII (2016–2020)

Sponsoring Agency:

Centers for Disease Control and Prevention (CDC), U.S. Department of Health and Human Services

Contact:

Division of Reproductive Health
National Center for Chronic Disease Prevention and Health Promotion
CDC
1600 Clifton Rd.
Atlanta, GA 30333
Telephone: 800-232-4636
<http://www.cdc.gov/PRAMS/>

Availability:

Information on obtaining data is available at <https://www.cdc.gov/prams/prams-data/researchers.htm> and by email to PRAMSProposals@cdc.gov. Data available for release by the CDC includes years 1988–2019.

Overview:

PRAMS is a CDC and state health department surveillance system. It collects state-specific, population-based data on maternal attitudes, behaviors, and experiences that occur several months before conception, during pregnancy, and immediately following delivery. The annual data sets include data from three sources: questionnaire data containing responses from mothers to the survey questionnaire; birth certificate data containing information on selected maternal characteristics (e.g., race, ethnicity, age) and pregnancy outcomes (e.g., birth weight, gestational age); and operations data generated by the PRAMS operational software, which include details about how the questionnaire was administered and are used primarily for operational evaluations and analyses of survey methods.

Survey Design/Methodology:

Each month, mothers who are state residents and have recently delivered a live-born infant during the preceding 2–4 months are randomly selected from a file of birth certificate records using stratified systematic sampling. Mothers who gave birth outside their state of residence and mothers who had a multiple birth greater than three gestations are excluded from the sampling frame. Selected mothers are mailed a questionnaire (the questionnaire is also available in Spanish), with telephone interview follow-up for nonrespondents. The PRAMS questionnaire has three parts: a core

that all states use; a bank of standardized optional questions that states may select from; and state-developed questions that are usually used only by the state that developed them. Some questions change with each revision or new phase of the questionnaire; however, most indicators can be compared across phases.

Sample Characteristics:

Forty-seven states, New York City, Puerto Rico, the District of Columbia, and the Great Plains Tribal Chairmen's Health Board currently participate in PRAMS Phase VIII, representing approximately 83 percent of all U.S. live births. Each participating state samples between 1,300 and 3,400 women per year.

Alcohol Variables:

The core questionnaire gathers data on the frequency of drinking and binge drinking before and during pregnancy and the education received from health care professionals regarding the effects of drinking on pregnancy.

Other Variables:

Substance use variables other than alcohol include tobacco, heroin, marijuana, and other drug use before and during pregnancy. In addition to demographic information, other variables include birth control usage, prenatal care, health problems during pregnancy, stressful life events during pregnancy, incidence of domestic violence, health of the newborn, breastfeeding practices, and use of health care and insurance.

The Study of Women's Health Across the Nation (SWAN)—1996–2008

Sponsoring Agency:

National Institute on Aging, the National Institute of Nursing Research, the National Institutes of Health, Office of Research on Women's Health, and the National Center for Complementary and Alternative Medicine

Contact:

SWAN Coordinating Center
University of Pittsburgh
4420 Bayard Street
615 Schenley Place
Pittsburgh, PA 15260
rs340348@pitt.edu
<https://www.swanstudy.org/>

Availability:

Requests for data access can be sent to the SWAN Coordinating Center at the University of Pittsburgh (email: swanaccess@edc.pitt.edu). Selected data from baseline up to the 10th follow-up visit are also available in a publicly accessible repository at <https://www.icpsr.umich.edu/icpsrweb/ICPSR/series/253>. The SWAN cross-sectional screener data set has been archived at the National Archive of Computerized Data on Aging. Further information regarding data access is at: <https://www.swanstudy.org/swan-research/data-access/>

Overview:

SWAN is a multisite longitudinal, epidemiologic study designed to examine the health of women during their middle years. The study examines the physical, biological, psychological, and social changes during this transitional period. The goal of SWAN's research is to help scientists, health care providers, and women learn how midlife experiences affect health and quality of life during aging.

In 1994, SWAN was designed as a multisite, observational study to be conducted in three phases. The initial phase consisted of focus groups of women with characteristics like those subsequently enrolled in SWAN. The second phase was a cross-sectional survey, a 15-minute computer-assisted telephone or in-person interview, conducted from 1995 through 1997. In the third phase, eligible premenopausal women were enrolled through seven designated research centers from the cross-sectional phase into the longitudinal follow-up study. Women who met the eligibility criteria were ages 42–52, had a uterus and at least one intact ovary, reported a menstrual period within the past three months, and had not taken hormone medications (such as birth control pills, estrogen or progesterone preparations) in the last three months. All participants underwent annual examinations that included interviews, anthropometry,

questionnaires, and a blood draw for the assessment of sociodemographic factors, cardiovascular disease risk factors, and reproductive hormone levels. SWAN has completed the screening, baseline, and 16 follow-up visits. Data collected from baseline through visit 10 are currently available to the public.

Survey Design/Methodology:

Enrollment into the longitudinal phase (Phase 3) began in January 1996. All seven clinical sites completed the baseline visit by December 1997. The annual visit included the following core components: physical measures (weight, height, hip, waist, and blood pressure), fasting morning blood draw, and interviewer-administered and self-administered questionnaires. Women were also given menstrual calendars to complete monthly over the next year.

Sample Characteristics:

The cross-sectional survey in the second phase was administered to 16,065 women who met the following criteria: residence within the geographic area specified by the clinical site; ability to speak and read English or the designated language (Chinese, Japanese, or Spanish) of the clinical site, ages 40–55; and self-identification within one of the two race/ethnicity groups studied at the clinical site. A total of 3,302 eligible women were enrolled into the longitudinal study population

and completed the baseline study (1,550 Caucasian, 935 African American, 286 Hispanic, 250 Chinese, and 281 Japanese).

Alcohol Variables:

The survey includes questions regarding current drinking, drinking since previous interview, consumption, and types of alcoholic beverages.

Other Variables:

Other variables include self-reported demographic factors (age, race/ethnicity, educational attainment, employment, marital status, number of children, ability to pay for basics), lifestyle factors (smoking, physical activity), height and weight, and menopausal status. The data also

include questions about doctor visits; medical conditions; medications; treatments; medical procedures; relationships; smoking; exposure to secondhand tobacco smoke; and menopause-related information such as age at pre-, peri- and post-menopause, self-attitudes, feelings, and common physical problems associated with menopause. A single in-person visit (visit 15) was included for the entire cohort as well as two bone mineral density (BMD) measurements for women participating in the BMD protocol (visit 16). At visit 15, SWAN participants completed new measures of physical function; physical activity; sleep; cognition; and vaginal, urogenital, and sexual health.

Survey of Inmates in State and Federal Correctional Facilities—1974, 1979, 1986, 1991, 1997, and 2004; Survey of Prison Inmates—2016

Sponsoring Agency:

Bureau of Justice Statistics, U.S. Department of Justice

Contact:

National Archive of Criminal Justice Data

P.O. Box 1248

Ann Arbor, MI 48106-1248

Telephone: 800-999-0960

nacjd@icpsr.umich.edu

<http://www.icpsr.umich.edu/icpsrweb/NACJD/>

Availability:

Data files are available for download from and information on accessing the restricted data can be found at <http://www.icpsr.umich.edu/icpsrweb/NACJD/series/70/studies?sortBy=7>.

Overview:

This survey is designed to provide nationally representative data on the characteristics of state prison inmates and sentenced federal inmates held in federally owned and operated facilities. The survey is conducted by the U.S. Census Bureau for the U.S. Department of Justice and collects information on current offenses and sentences; criminal history; family background and personal characteristics; prior drug and alcohol use and treatment programs; gun possession and use; and prison activities, programs, and services. Prior surveys of state prison inmates, the Survey of Inmates of State Correctional Facilities, were conducted in 1974, 1979, 1986, and 1991. Sentenced federal prison inmates were first interviewed in 1991, and the federal data are combined with the state data in the 1991 and 1997 surveys. The survey was renamed the Survey of Prison Inmates with the 2016 implementation.

Survey Design/Methodology:

The survey used a stratified, two-stage selection process. In the first stage, correctional facilities were separated into two sampling frames, and a systematic sample of facilities was selected within strata on each frame with probabilities proportional to the size of each facility. In the second stage, interviewers visited each selected facility and systematically selected a sample of male and female inmates using predetermined procedures.

Sample Characteristics:

The survey sample is selected from correctional facilities using a stratified, two-stage selection divided into male/female facilities, census region, and facility type. The 1974 survey included about 10,000 inmates, and the 1979 and 1986 surveys included 11,397 and 13,711 inmates, respectively. The 1991 survey included a total of 20,558 inmates from 277 prisons and 53 federal facilities, and the 1997 survey included a total of 18,326 inmates. The 2016 survey included 24,848 inmates from 306 prisons and 58 federal facilities.

Alcohol Variables:

Alcohol variables include the following: overall frequency of drinking in the year before arrest, whether drinking occurs on a regular basis, age when first began drinking regularly, self-perception of degree of drunkenness reached at end of a typical drinking session, and treatment history.

Other Variables:

Substance use variables other than alcohol include tobacco, heroin, marijuana, and other drugs. Other variables include age, sex, race/ethnicity, marital status, education, family background, income in year before offense, employment in year before offense, current offense, number of prior convictions, drug-related crime, gang membership, use of weapons, and needle sharing. Data are also collected on military service, prison activities, and involvement in programs and services.

Youth Risk Behavior Survey (YRBS)—1991–2019 (High School), Biennially, 1998 (Alternative High School), 1995 (College), and 1992 (NHIS)

Sponsoring Agency:

Division of Adolescent and School Health, National Center for Chronic Disease Prevention and Health Promotion, Centers for Disease Control and Prevention (CDC), U.S. Department of Health and Human Services

Contact:

Division of Adolescent and School Health
National Center for HIV/AIDS, Viral Hepatitis, STD, and TB Prevention
CDC
4770 Buford Highway, NE, Mail Stop K-33
Atlanta, GA 30341-3717
Telephone: 800-232-4636
<https://www.cdc.gov/healthyouth/data/yrbs/index.htm>

Availability:

Data files are available for download from <https://www.cdc.gov/healthyouth/data/yrbs/data.htm>.

Overview:

The CDC established the Youth Risk Behavior Surveillance System (YRBSS) to monitor health-risk behaviors among youth and to assess trends in such behaviors over time; YRBS is a component. YRBS measures youth risk behaviors in six risk areas: (1) behaviors that contribute to unintentional injuries and violence, (2) tobacco use, (3) alcohol and other drug use, (4) sexual behaviors that contribute to unintended pregnancy and sexually transmitted diseases including HIV infection, (5) unhealthy dietary behaviors, and (6) physical inactivity. Since 1991, data have been collected biennially, and the latest available survey data are for 2015. The 1992 YRBS was a supplement of the 1992 NHIS (see page 42). The 1998 National Alternative High School Youth Risk Behavior Survey was conducted to measure selected health-risk behaviors among a nationally representative sample of students in grades 9–12 attending alternative high schools.

Survey Design/Methodology:

YRBSS includes national, state, territorial, Tribal government, and local school-based surveys using representative samples of 9th through 12th grade students. The CDC conducts the national survey; the additional surveys are conducted by departments of health and education and are representative of mostly public high school students in each jurisdiction. The national YRBS uses a three-stage cluster sample design to produce a nationally representative sample of U.S. high school students. The sampling frame includes primary sampling units (PSUs)

consisting of large-sized counties or groups of smaller, adjacent counties from which schools are selected. One or two entire classes in each chosen school and in each of grades 9–12 are then randomly selected. The state, territorial, Tribal government, and local surveys employ two-stage cluster design in which schools are first selected with probability proportional to school enrollment size. In the second sampling stage, intact classes of a required subject or period are selected randomly. All YRBSS questionnaires are self-administered, and students record their responses on a computer-scannable answer sheet.

Sample Characteristics:

YRBS uses national, school-based samples of 11,000 to 16,000 students in the 9th through 12th grades. For the 2019 national YRBS, 13,872 questionnaires were completed in 136 public and private schools. Black and Hispanic high school students were oversampled. YRBS is not designed to represent individual states, so performing state-level analyses is not recommended.

Alcohol Variables:

Alcohol questions include age at first drink, lifetime drinking, frequency and quantity of drinking, and alcohol use prior to sexual intercourse within the past 30 days are also on school property within the past 30 days was measured. A variable was added in 2007 that asks participants how they usually obtained the alcohol they drank in the last 30 days.

Other Variables:

Other variables include seatbelt and helmet use; physical fighting and carrying weapons; suicide

attempts; tobacco use; use of marijuana, cocaine, steroids, or other illegal drugs; HIV awareness; sexual activity; diet; and physical activity.

Section 3:

AEDS Publications and Products

Below is the list of AEDS-produced publications, based on epidemiologic research.

Data Reference Manuals

This series of manuals provides extensive coverage of data on alcohol consumption, alcohol use disorders, drinking patterns, drinking-related risk behavior, and alcohol-related morbidity and mortality.

U.S. Alcohol Epidemiologic Data Reference Manual, Volume 10, Alcohol Use and Alcohol Use Disorders in the United States: Main Findings from the 2012–2013 National Epidemiologic Survey on Alcohol and Related Conditions-III (NESARC-III). April 2016. NIH Publication No. 16-AA-8020.

U.S. Alcohol Epidemiologic Data Reference Manual, Volume 9, Alcohol-Related Emergency Department Visits and Hospitalizations and Their Co-Occurring Drug-Related, Mental Health, and Injury Conditions in the United States: Findings from the 2006–2010 National Emergency Department Sample (NEDS) and Nationwide Inpatient Sample (NIS). September 2013. NIH Publication No. 13-8000.

U.S. Alcohol Epidemiologic Data Reference Manual, Volume 8, Number 2, Alcohol Use and Alcohol Use Disorders in the United States, A 3-Year Follow-Up: Main Findings from the 2004–2005 Wave 2 National Epidemiologic Survey on Alcohol and Related Conditions (NESARC). September 2010. NIH Publication No. 10-7677.

U.S. Alcohol Epidemiologic Data Reference Manual, Volume 8, Number 1, Alcohol Use and Alcohol Use Disorders in the United States: Main Findings from the 2001–2002 National Epidemiologic Survey on Alcohol and Related Conditions (NESARC). January 2006. NIH Publication No. 05-5737.

U.S. Alcohol Epidemiologic Data Reference Manual. Volume 1, Fourth Edition: U.S. Apparent Consumption of Alcoholic Beverages Based on State Sales, Taxation, or Receipt Data. June 2004. NIH Publication No. 04-5563.

U.S. Alcohol Epidemiologic Data Reference Manual. Volume 7, State Trends in Drinking Behaviors, 1984–2001. June 2003. NIH Publication No. 02-5213.

Alcohol Consumption and Problems in the General Population: Findings from the 1992 National Longitudinal Alcohol Epidemiologic Survey. June 2002. NIH Publication No. 02-4997.

U.S. Alcohol Epidemiologic Data Reference Manual. Volume 6, First Edition: Drinking in the United States: Main Findings from the 1992 National Longitudinal Alcohol Epidemiologic Survey (NLAES). November 1998. NIH Publication No. 99-3519.

U.S. Alcohol Epidemiologic Data Reference Manual. Volume 5, First Edition: State Trends in Alcohol-Related Mortality, 1979–92. September 1996. NIH Publication No. 96-4174.

U.S. Alcohol Epidemiologic Data Reference Manual. Volume 3, Fourth Edition: County Alcohol Problem Indicators, 1986–1990. July 1994. NIH Publication No. 94-3747.

Electronic copies of these manuals are available on the NIAAA website:

<http://pubs.niaaa.nih.gov/publications/manual.htm>. Hard copies of NIH Publication No. 02-4997 and manual volumes 1, 3, 5, 7, 8, 9, and 10 may be ordered online.

AEDS Surveillance Reports

AEDS prepares surveillance reports that monitor long-term trends in alcohol use and its consequences. Surveillance topics include per capita alcohol consumption, liver cirrhosis mortality, underage drinking, hospital discharges for alcohol-related conditions, substance use among reproductive age females, and alcohol-related traffic crashes. The most current issues are listed below:

Surveillance Report #119: Apparent Per Capita Alcohol Consumption: National, State, and Regional Trends, 1977–2020. Slater, M.E.; Alpert, H.R. National Institute on Alcohol Abuse and Alcoholism, Division of Epidemiology and Prevention Research, Alcohol Epidemiologic Data System, April 2022.

Surveillance Report #118: Liver Cirrhosis Mortality in the United States: National, State, and Regional Trends, 2000–2019. Chen, C.M.; Yoon, Y-H. National Institute on Alcohol Abuse and Alcoholism, Division of Epidemiology and Prevention Research, Alcohol Epidemiologic Data System, February 2022.

Surveillance Report: COVID-19: Alcohol Sales During the COVID-19 Pandemic. National Institute on Alcohol Abuse and Alcoholism, Division of Epidemiology and Prevention Research, Alcohol Epidemiologic Data System, <https://pubs.niaaa.nih.gov/publications/surveillance-covid-19/COVSALES.htm>.

Surveillance Report #116: Trends in Underage Drinking in the United States, 1991–2019. Chen, C.M.; Yoon, Y.H. National Institute on Alcohol Abuse and Alcoholism, Division of Epidemiology and Prevention Research, Alcohol Epidemiologic Data System, March, 2021.

Surveillance Report #112: Trends in Alcohol-Related Morbidity Among Community Hospital Discharges, United States, 2000–2015. Chen, C.M.; Yoon, Y-H. National Institute on Alcohol Abuse and Alcoholism, Division of Epidemiology and Prevention Research, Alcohol Epidemiologic Data System, August 2018.

Surveillance Report #109: Trends in Substance Use Among Reproductive-Age Females in the United States, 2002–2015. Slater, M.E.; Haughwout, S.P. National Institute on Alcohol Abuse and Alcoholism, Division of Epidemiology and Prevention Research, Alcohol Epidemiologic Data System, September 2017.

Surveillance Report #76: Trends in Alcohol-Related Fatal Traffic Crashes, United States, 1977–2004. Yi, H.; Chen, C.M.; Williams, G.D. National Institute on Alcohol Abuse and Alcoholism, Division of Epidemiology and Prevention Research, Alcohol Epidemiologic Data System, August 2006.

Full text is available on the NIAAA website: <http://pubs.niaaa.nih.gov/publications/surveillance.htm>.

APPENDIX: LIST OF ACRONYMS

| | |
|-------------------|---|
| ABES | Adolescent Behaviors and Experiences Survey |
| ACHA | American College Health Association |
| ADAM | Arrestee Drug Abuse Monitoring |
| AEDS | Alcohol Epidemiologic Data System |
| Add Health | National Longitudinal Study on Adolescent to Adult Health |
| AHRQ | Agency for Health Care Research and Quality |
| AUD | Alcohol Use Disorder |
| AUDIT | Alcohol Use Disorders Identification Test |
| BAC | Blood Alcohol Concentration |
| BJS | Bureau of Justice Statistics |
| BRFSS | Behavioral Risk Factors Surveillance System |
| CARDIA | Coronary Artery Risk Development in Young Adults |
| CDC | Centers for Disease Control and Prevention |
| CIFSAD | Collaborative Initiative on Fetal Alcohol Spectrum Disorder |
| CIRP | Cooperative Institutional Research Program |
| COGA | Collaborative Studies on Genetics of Alcoholism |
| COVID-19 | Coronavirus Disease of 2019 |
| CRSS | Crash Report Sampling System |
| CSS | College Senior Survey |
| CVS | Cardiovascular Disease |
| DAWN | Drug Abuse Warning Network |
| DOT | Department of Transportation |
| DRG | Diagnostic Related Groups |
| DSM | Diagnostic and Statistical Manual |
| DWI | Driving While Intoxicated |
| DUI | Driving Under the Influence |
| EHR | Electronic Health Record |
| EMS | Emergency Medical Services |
| FARS | Fatality Analysis Reporting System (formerly Fatal Accident Reporting System) |
| GES | General Estimates System |
| HCHS/SOL | Hispanic Community Health Study / Study of Latinos |
| HCUP | Health Care Cost and Utilization Project |
| HINTS | Health Information National Trends Survey |
| HMS | Healthy Minds Study |
| HPDP | Health Prevention Disease Promotion Supplement of NHIS |
| HHS | Department of Health and Human Services |
| HRS | Health and Retirement Study: A Longitudinal Study of Health, Retirement, and Aging |
| ICD-9 | International Classification of Diseases, Ninth Revision [mortality] |
| ICD-9-CM | International Classification of Diseases, Ninth Revision, Clinical Modification [morbidity] |
| ICD-10 | International Classification of Diseases, Tenth Revision [mortality] |
| ICPSR | Inter-university Consortium for Political and Social Research |
| KID | Kids' Inpatient Database |

| | |
|-----------------|--|
| MCD | Multiple Cause of Death |
| MIDUS | Midlife in the United States |
| MTF | Monitoring the Future |
| N3C | National COVID Cohort Collaborative Data Enclave |
| NACJD | National Archive of Criminal Justice Data |
| NAMCS | National Ambulatory Medical Care Survey |
| NAS | National Alcohol Survey |
| NASS | National Ambulatory Surgery Sample |
| NATSCEV | National Survey of Children's Exposure to Violence |
| NCANDS | National Child Abuse and Neglect Data System |
| NCATS | National Center for Advancing Translational Sciences |
| NCHS | National Center for Health Statistics |
| NCVS | National Crime Victimization Survey |
| NDATUS | National Drug and Alcoholism Treatment Unit Survey |
| NEMESIS | National Emergency Medical Services Information System |
| NESARC | National Epidemiologic Survey on Alcohol and Related Conditions |
| NHAMCS | National Hospital Ambulatory Medical Care Survey |
| NHANES | National Health and Nutrition Examination Survey |
| NHCA | National College Health Assessment |
| NHCS | National Hospital Care Survey |
| NHDS | National Hospital Discharge Survey |
| NHIS | National Health Interview Survey |
| NHLBI | National Heart Lung and Blood Institute |
| NHS | Nurses' Health Study |
| NHTSA | National Highway Traffic Safety Administration |
| NIAAA | National Institute on Alcohol Abuse and Alcoholism |
| NICHD | National Institute of Child Health and Human Development |
| NIDA | National Institute on Drug Abuse |
| NIH | National Institutes of Health |
| NIMH | National Institute of Mental Health |
| NIS | National Inpatient Sample |
| NLSCYA | National Longitudinal Survey of Youth 1979 Child and Young Adult |
| NLSY | National Longitudinal Survey of Youth |
| N-MHSS | National Mental Health Services Survey |
| NRD | National Readmissions Database |
| NSDUH | National Survey of Drug Use and Health |
| NSFG | National Survey of Family Growth |
| N-SSATS | National Survey of Substance Abuse Treatment Services |
| N-SUMHSS | National Substance Use and Mental Health Services Survey |
| NSYC | National Survey on Youth in Custody |
| NVDRS | National Violent Death Reporting System |
| PATH | Population Assessment of Tobacco and Health |
| PRAMS | Pregnancy Risk Assessment Monitoring System |

| | |
|---------------|--|
| PSID | Panel Study of Income Dynamics |
| PSU | Primary Sampling Unit |
| QF | Quantity–Frequency |
| RANDS | Research and Development Survey |
| SAMHDA | Substance Abuse and Mental Health Data Archive |
| SAMHSA | Substance Abuse and Mental Health Services Administration |
| SID | State Inpatient Databases |
| SMSA | Standard Metropolitan Statistical Area |
| SOLNAS | Study of Latinos: Nutrition and Physical Activity Assessment Study |
| SWAN | Study of Women’s Health Across the Nation |
| TEDS | Treatment Episode Data Set |
| TFS | CIRP Freshman Survey |
| UAS | Understanding America Study |
| UFDS | Uniform Facility Data Set |
| YRBS | Youth Risk Behavior Survey |
| YRBSS | Youth Risk Behavior Surveillance System |