Published in final edited form as:

J Am Coll Health. 2016 April ; 64(3): 262–267. doi:10.1080/07448481.2015.1088018.

Reaching a representative sample of college students: A comparative analysis

Daniel P. Giovenco, MPH,

Center for Tobacco Studies, Rutgers School of Public Health, 732-235-9338

Daniel A. Gundersen, PhD, and

Department of Family Medicine and Community Health, Rutgers University Robert Wood, Johnson Medical School

Cristine D. Delnevo, PhD, MPH

Center for Tobacco Studies, Rutgers School of Public Health

Daniel P. Giovenco: d.giovenco@rutgers.edu

Abstract

Objective—To explore the feasibility of a random-digit dial (RDD) cellular phone survey in order to study college health.

Methods—Demographic distributions from the 2011 National Young Adult Health Survey (NYAHS) were benchmarked against enrollment numbers from the Integrated Postsecondary Education System (IPEDS). The sample quality was compared to the Harvard School of Public Health College Alcohol Study (1993), National College Health Risk Behavior Survey (1995), and National College Health Assessment (2011).

Results—Overall, the NYAHS performed as well, if not better, than previous college health surveys at reaching important demographic sub-groups.

Conclusions—Cellular phone RDD is a feasible approach for reaching college students and offers a simpler alternative to traditional complex sampling designs.

According to the U.S. Census, over 23 million adults in the United States are currently enrolled in a college or university.¹ Moreover, undergraduate enrollment increased in the past decade, and rates are expected to grow.² Although full-time, undergraduate students at 4-year institutions are the single largest group of college enrollees, students in part-time or 2-year programs make up a sizable proportion of this demographic group.² Monitoring college students' health behaviors is important; the undergraduate college experience often marks the transition into adulthood and is frequently associated with risky behaviors, such as binge drinking and use of other drugs.^{3,4} Additionally, health issues facing “non-traditional” students (e.g., older students, those in community colleges, or with part-time status) are understudied but may differ from the health issues of more “typical” college students.⁵ For example, nearly a third of community college students work full-time and have at least one dependent, which may influence their physical or mental health.⁶

There is a growing need for ongoing research focused on college students' health behaviors. Reaching a nationally representative sample of college students, however, is operationally

challenging and requires considerable resources. To date, only two health surveys have been conducted where a nationally representative sample of college students in the U.S. was recruited. First, the Harvard School of Public Health's College Alcohol Study (CAS) administered its inaugural survey in 1993, collecting data via mail survey from over 15,000 students at 179 institutions.⁷ Although the CAS provided a detailed look at high-risk behaviors, the sample was limited to full-time undergraduate students and did not capture the behaviors of "non-traditional" college students. Second, the Centers for Disease Control and Prevention (CDC) expanded its Youth Risk Behavior Surveillance System (YRBS) to include the National College Health Risk Behavior Survey (NCHRBS) in 1995.⁸ Despite the survey's utility documenting a wide variety of health behaviors, the NCHRBS has not been re-administered, likely due to its high complexity and the operational burden of coordinating a mail survey with approximately 150 institutions.

A third survey, the National College Health Assessment (NCHA), bears mentioning because of its large scope and wide use. Managed by the American College Health Association (ACHA), it is currently the most widely-used tool for assessing the health behaviors of college students. While the NCHA has broad reach (there were 129 institutions and over 105,000 respondents in Spring 2011), the schools are self-selecting, thereby limiting generalizability. The ACHA recommends that the sample be used as a "reference group" to which others can compare data.⁹ Despite a non-probability sampling design, the NCHA produced similar estimates of health behaviors when evaluated against the CAS and NCHRBS.⁹

The surveys described above have substantial operational challenges and require considerable resources to coordinate data collection across multiple institutions. Consequently, many college student surveys have been limited to single universities¹⁰⁻¹³ or a collection of universities^{14,15} that do not produce nationally representative findings. Large-scale, national population health surveys like the Behavioral Risk Factor Surveillance System (BRFSS) have the potential to capture rich health data on college students with less operational difficulty. As presently implemented, however, they may not sufficiently reach this population. The BRFSS has historically relied on random-digit dialing (RDD) of landline telephone exchanges, which excluded those living in dormitories.¹⁶ In 2011, cellular (cell) phones were added to the sampling frame, likely increasing coverage of those in college. Contrary to landlines, cell phones are "attached" to a person, meaning they can potentially reach students regardless of domicile. At present, however, the BRFSS¹⁷ does not assess current enrollment in a college or university.

College health researchers would benefit from a sampling approach with potentially fewer operational barriers to adequately reach a national sample of college students. While telephone-based surveys such as the BRFSS formerly had little methodological value in surveys of the college-aged population, the recent growth in cell phone ownership gives merit to RDD of cell phones as a valid sampling technique. The National Young Adult Health Survey (NYAHS), first administered in 2011, assessed the feasibility of using RDD of cell phones to reach a representative sample of young adults between the ages of 18-34. The unweighted sample was demographically similar to the U.S. Census, a rarity for population telephone surveys, which tend to be skewed toward females, non-Hispanic

Whites, and older adults.¹⁸ To evaluate the sample quality of RDD of cell phones to study college health, this comparative analysis benchmarks the demographic profile of the NYAHS against college enrollment numbers from the Integrated Postsecondary Education System (IPEDS) and compares its quality to the CAS, NCHRBS, and NCHA.

Methods

National Young Adult Health Survey

The NYAHS is a national cell phone RDD survey of young adults between the ages of 18-34. Participants' telephone numbers were randomly selected from cellular-dedicated thousand-level blocks (NPA-NXX-Z000 to NPA-NXX-Z999) stratified by the 4 U.S. Census regions, originating from the Telcordia Local Exchange Routing Guide (iconectiv, Piscataway, New Jersey). Interviewers used a screening question to identify eligible participants between the ages of 18 and 34. Data were collected via computer-assisted telephone interviewing (CATI) by ICF International, an independent research firm with extensive survey research experience.¹⁹ ICF trained all interviewers over a 3-day period and monitored live interviews during data collection to ensure accuracy of question wording and data entry. Although the instrument was primarily designed to collect data related to tobacco use, it also included questions about other health behaviors. After each interview, participants received a \$10 e-gift card to offset any cellular charges incurred. Of the 2,871 respondents in the NYAHS, 1,193 indicated they were college students. Detailed descriptions of the NYAHS methodology have been published elsewhere.^{18,20} The Rutgers Biomedical and Health Sciences Institutional Review Board approved all procedures.

The NYAHS response rate was 24% with a cooperation rate of 64.2%. Although the response rate appears low, this is largely a function of high rates of “undetermined eligibility” among cell phones, resulting from cellular numbers assigned to devices like computers, tablets, and fax machines, and the difficulty identifying business telephones. These factors make traditional response rate calculations for cell phone surveys problematic.²¹ Indeed, the NYAHS response rate was comparable to the BRFSS cell phone response rate.^{18,22} Moreover, the sample closely matched the census distribution for young adults, even more so than the dual-frame (i.e., cell phone and landline) young adult sample in the BRFSS.

Other College-Health Surveys

Two surveys in this comparative analysis, the CAS and the NCHRBS, used a two-stage cluster sampling design, selecting institutions based on probability proportionate to enrollment size.^{7,8} The CAS oversampled women's colleges and institutions with fewer than 1,000 students, and the NCHRBS oversampled Black and Hispanic students.^{7,8} The NCHA is a convenience sample of institutions that pay to participate. Only schools that randomly sample students are included in the national dataset.⁹ The methods of the NYAHS and the college surveys used as sources of comparison in this analysis are summarized in Table 1.

Gold Standard

Conducted annually by the National Center for Education Statistics (NCES), the IPEDS is designed to collect enrollment data from every postsecondary institution that participates in or are applicants for any federal student assistance program authorized by Title IV of the Higher Education Act of 1965. Over 7,500 institutions are required to participate, including colleges, universities, community colleges, and technical schools.²³

Statistical Analysis

To determine how closely each survey sample resembled the demographic characteristics of U.S. college students, unweighted sex, age-group, and race/ethnicity distributions were compared to IPEDS data from the respective survey years. The unweighted comparison examines how well the sampling methodology reached important population subgroups. As a summary measure of sample quality, the mean absolute deviations across key demographics were calculated for each survey using the IPEDS as the source of population distributions. For this analysis, the NYAHS sample was restricted to the 1,193 individuals reporting that they were enrolled in a college or university in the past six months.

Results

Table 2 displays the demographics for each survey along with year-specific IPEDS demographics. Females make up just over half of U.S. college students, but they were slightly overrepresented in NCHRBS and NCHA. In those surveys, the proportion of females deviated from IPEDS enrollment percentages by 6.3 and 8.4 points, respectively. The NYAHS was the only survey in which the proportion of males exceeded the national estimate.

Generally, the NCHRBS and NYAHS best approximated the actual age distribution of college students; specifically, these surveys captured students over age 25 better than the CAS and NCHA. It should be noted that the CAS surveyed only full-time undergraduate students, who tend to be younger in age than “non-traditional” students. Additionally, all surveys but the CAS restricted participation to students 18 years and older, and so did not reach students under 18.

The NCHRBS oversampled institutions with higher percentages of Black and Hispanic students; unsurprisingly, they reached these groups better than the other surveys in this analysis. The NCHA deviated greatly from national estimates; almost three-quarters of respondents reported being White, an absolute deviation of 13.5 percentage points from the IPEDS.

Overall, the NYAHS had the lowest mean absolute deviation from IPEDS enrollment proportions, at 3.2 percent. The NCHRBS also closely approximated demographic benchmarks, deviating 3.8 percentage points, on average. The CAS and NCHA differ most greatly from the demographic distribution of the IPEDS; this is largely driven by the underrepresentation of older students, and in the case of the NCHA, the overrepresentation of White students.

Comment

Current and generalizable estimates of health behaviors among college students are often unavailable and not continually updated, primarily due to the complexity of recruiting representative samples of students. Surveys like the CAS, NCHRBS, and NCHA have played essential roles in assessing the health of the college-aged population, but each has a unique purpose and distinct limitations. A simpler surveying method that ensures high sample quality and enables researchers to customize survey instruments would greatly benefit this field of research.

Although the NYAHS was not specifically designed to reach college students, it demonstrated the utility of cell phone RDD in recruiting this population efficiently. With just 1,193 respondents, this method most closely approximated the actual demographic distribution of college students in the U.S. Other surveys, like the CAS and NCHRBS, adequately reached various subgroups, but likely had high costs and administrative complexity from having to coordinate with over 100 geographically-dispersed institutions. The 2-stage cluster sampling design requires inclusion of a large number of colleges, which helps reduce the sample design effects and ensures an effective sample size with adequate statistical power. The NCHA, though useful to document the health issues facing college students, deviated substantially from national estimates of college enrollment demographics, potentially introducing biased estimates of health behaviors.

Limitations

This comparative analysis is subject to limitations. To our knowledge, the CAS and the NCHRBS were the only two health-related surveys that successfully reached a nationally representative sample of college students, but it is possible that we may have missed other, smaller efforts. The ACHA uses these two datasets to evaluate the reliability and validity of the NCHA, so we felt that they were appropriate to use in our analysis. Second, IPEDS is by no means a perfect estimate of college enrollment in the U.S. Indeed, not every postsecondary institution participates in federal student financial aid programs, but since all Title IV degree-granting institutions are included in the database, as well as thousands of other institutions, it is widely considered an accurate representation of postsecondary education trends.²³ The NCHA was the only survey that allowed multiple responses for the racial/ethnic identity question. While this may partially explain the elevated proportion of White students, the underrepresentation of Blacks and Hispanics is still a concern. Finally, although the NYAHS included a question assessing current college enrollment status, it did not collect information about the type of college, its location, and other institutional features.

Conclusion

Each survey in the present analysis served a unique purpose and generated important data that fulfilled the needs of the investigators and/or the participating institutions. The intent of this study was not to discredit any particular survey, but rather to evaluate the sample quality of a cell phone RDD design relative to surveys with more common sampling strategies for this population. Moreover, it offers a new and potentially simpler strategy that may be more effective in reaching all types of young adults enrolled in colleges and universities. Although

cell phone surveys are not inexpensive (the final “cost per complete” rate of the NYAHS was \$158, which includes set up and fielding costs, sample selection, and participant incentives), they eliminate the need for coordination between hundreds of institutions and can produce generalizable estimates with smaller sample sizes than traditional college surveying methods. As a result, cell phone surveys can potentially be administered more rapidly and regularly than traditional methods, an important advantage when assessing emergent health behaviors among college students. Given the relative novelty of cell phone RDD techniques for health-based surveys, this methodology warrants further study among college-aged students. Future research should assess how well cell phone RDD surveys reach different types of college students (e.g. 2-year, 4-year, full-time, part-time) and compare these measures to existing college datasets.

References

1. U.S. Census Bureau. [Accessed January 28, 2014] School enrollment. 2013 American Community Survey 1-year estimates - Table S1401. Available at: http://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid=ACS_13_1YR_S1401&prodType=table
2. National Center for Education Statistics; U.S. Department of Education. [Accessed January 28, 2014] Digest of education statistics: 2012. Available at: http://nces.ed.gov/programs/digest/d12/ch_3.asp
3. Johnston, LD.; O'Malley, PM.; Bachman, JG.; Schulenberg, JE.; Miech, RA. [Accessed January 28, 2015] Monitoring the Future national survey results on drug use, 1975–2013: Volume 2. College students and adults ages 19–55. Available at: http://www.monitoringthefuture.org/pubs/monographs/mtf-vol2_
4. Wechsler H, Nelson TF. What we have learned from the Harvard School of Public Health College Alcohol Study: Focusing attention on college student alcohol consumption and the environmental conditions that promote it. *J Stud Alcohol Drugs*. 2008; 69(4):481–490. [PubMed: 18612562]
5. Pokhrel P, Little MA, Herzog TA. Current methods in health behavior research among U.S. community college students: A review of the literature. *Eval Health Prof*. 2013; 37(2):178–202.10.1177/0163278713512125 [PubMed: 24227658]
6. Juskiewicz, J. [Accessed January 28, 2015] Community college students and federal student financial aid: A primer. Available at: http://www.aacc.nche.edu/Publications/Reports/Documents/CCStudents_A_Primer.pdf
7. Wechsler H, Davenport A, Dowdall G, Moeykens B, Castillo S. Health and behavioral consequences of binge drinking in college: A national survey of students at 140 campuses. *JAMA*. 1994; 272(21):1672–1677. [PubMed: 7966895]
8. Centers for Disease Control and Prevention. Youth Risk Behavior Surveillance: National College Health Risk Behavior Survey - United States, 1995. *MMWR*. 1997; 46(SS-6):1–56.
9. American College Health Association. [Accessed January 29, 2015] About ACHA-NCHA: Generalizability, reliability, and validity analysis. Available at: <http://www.acha-ncha.org/grvanalysis.html>
10. DeSantis AD, Webb EM, Noar SM. Illicit use of prescription ADHD medications on a college campus: A multimethodological approach. *J Am Coll Health*. 2008; 57(3):315–323.10.3200/JACH.57.3.315-324 [PubMed: 18980888]
11. Gaultney JF. The prevalence of sleep disorders in college students: Impact on academic performance. *Am J Coll Health*. 2010; 59(2):91–97.10.1080/07448481.2010.483708
12. Racette SB, Deusinger SS, Strube MJ, Highstein GR, Deusinger RH. Weight changes, exercise, and dietary patterns during freshman and sophomore years of college. *J Am Coll Health*. 2005; 53(6):245–251.10.3200/JACH.53.6.245-251 [PubMed: 15900988]
13. Eisenberg D, Nicklett EJ, Roeder K, Kirz NE. Eating disorder symptoms among college students: Prevalence, persistence, correlates, and treatment-seeking. *J Am Coll Health*. 2011; 59(8):700–707.10.1080/07448481.2010.546461 [PubMed: 21950250]

14. Glover ED, Laflin M, Edwards SW. Age of initiation and switching patterns between smokeless tobacco and cigarettes among college students in the United States. *Am J Public Health*. 1989; 79(2):207–208.10.2105/AJPH.79.2.207 [PubMed: 2913845]
15. Laska MN, VanKim NA, Erickson DJ, Lust K, Eisenberg ME, Rosser BRS. Disparities in weight and weight behaviors by sexual orientation in college students. *Am J Public Health*. 2015; 105(1): 111–121.10.2105/AJPH.2014.302094 [PubMed: 25393177]
16. Centers for Disease Control and Prevention. [Accessed January 28, 2015] About the Behavioral Risk Factor Surveillance System (BRFSS). Available at: http://www.cdc.gov/brfss/about/about_brfss.htm
17. Pierannunzi C, Town M, Garvin W, Shaw FE, Balluz L. Methodologic changes in the Behavioral Risk Factor Surveillance System in 2011 and potential effects on prevalence estimates. *MMWR*. 2012; 61(22):410–413. [PubMed: 22672976]
18. Gundersen DA, ZuWallack RS, Dayton J, Echeverria SE, Delnevo CD. Assessing the feasibility and sample quality of a national random-digit dialing cellular phone survey of young adults. *Am J Epidemiol*. 2014; 179(1):39–47.10.1093/aje/kwt226 [PubMed: 24100957]
19. [Accessed July 13, 2015] Survey Research. ICF International Web site. <http://www.icfi.com/services/survey-research>
20. Gundersen DA, Peters KR, Conner A, Dayton J, Delnevo CD. Stability of sample quality for a national random-digit dialing cellular phone survey of young adults. *Am J Epidemiol*. 2014; 180(10):1047–1049.10.1093/aje/kwu271 [PubMed: 25352574]
21. American Association for Public Opinion Research Task Force. [Accessed February 2, 2015] New considerations for survey researchers when planning and conducting RDD telephone surveys in the U.S. with respondents reached via cell phone numbers. Available at: http://www.aapor.org/AAPORKentico/AAPOR_Main/media/MainSiteFiles/2010AAPORCellPhoneTFReport.pdf
22. Centers for Disease Control and Prevention. [Accessed February 6, 2015] Behavioral Risk Factor Surveillance System 2011: Summary data quality report - Version #5. Available at: http://www.cdc.gov/brfss/pdf/2011_Summary_Data_Quality_Report.pdf
23. National Center for Education Statistics; U.S. Department of Education. [Accessed February 2, 2015] Integrated Postsecondary Education Data System: About IPEDS. Available at: <http://nces.ed.gov/ipeds/about/>

Table 1
Summary of the CAS, NCHRBS, NCHA, and NYAHS sampling methodologies

Survey	Population	RR ^a	Methods
Harvard College Alcohol Study (1993)	Full-time, undergraduate students at 4-year institutions	69%	<ul style="list-style-type: none"> • 2-stage cluster sample • Stage 1: Random sample of 179 accredited 4-year universities (selected based on probability proportionate to enrollment size) • Oversampled small colleges and women-only colleges • Stage 2: Simple random sample of students • 4 mailings sent over approximately 30 days (two questionnaires and two reminder postcards) • Incentive was a cash prize drawing
National College Health Risk Behavior Survey (1995)	Full-time or part-time undergraduate students over 18 years old at 2- and 4-year institutions	60%	<ul style="list-style-type: none"> • 2 stage cluster sample • Stage 1: Random sample of 148 colleges and universities offering at least a 2-year degree (selected based on probability proportionate to enrollment size) • Oversampled schools with high Black & Hispanic enrollment • Stage 2: Simple random sample of students • Questionnaire sent by mail with no incentives
National College Health Assessment (Spring 2011)	2- and 4-year college students over 18	29%	<ul style="list-style-type: none"> • Institutions contact ACHA (i.e., the sample is self-selecting) and customize their surveying strategy (e.g., web-based or paper and pencil, use of incentives, reminder letters, custom questions, etc.) • Only schools that randomly sampled students or classrooms are included in the final "reference group" • 129 participating institutions
National Young Adult Health Survey (2011)	Adults between 18-34 years old	24.4%	<ul style="list-style-type: none"> • Randomly selected cell phone numbers stratified by U.S. Census region • Data collected via computer-assisted telephone interviewing • Incentive was a \$10 electronic gift card

^aRR: Response Rate

Table 2
Unweighted demographic distributions of the CAS, NCHRBBS, NCHA, and NYAHS and their deviations from IPEDS

	CAS ^a (1993) <i>n</i> =15,282	IPEDS ^b (1993)	NCHRBBS ^c (1995) <i>n</i> =4,814	IPEDS (1995)	NCHA ^d (2011) <i>n</i> =105,781	IPEDS (2011)	NYAHS ^e (2011) <i>n</i> =1,193	IPEDS (2011)
Sex								
Female	57.2	55.1	61.8	55.5	65.4	57.0	54.7	57.0
Male	42.9	44.9	38.2	44.5	34.6	43.0	45.4	43.0
Age Group^f								
Under 18	0.3	0.9	0.0	1.0	0.0	3.4	0.0	3.4
18-19	27.7	19.9	20.2	20.3	28.1	20.4	19.1	20.4
20-21	36.2	18.7	23.8	19.0	32.4	19.8	21.9	19.8
22-24	22.7	18.0	17.7	16.9	18.6	16.4	25.2	16.4
25+	13.2	42.6	38.3	42.8	20.8	40.0	33.9	40.0
Race/Ethnicity^g								
White, NH	79.3	74.1	62.5	72.3	72.5	59.0	58.9	59.0
Black, NH	5.8	9.9	13.7	10.3	5.8	14.6	11.7	14.6
Hispanic	6.4	6.9	14.9	7.7	7.8	13.8	11.0	13.8
Asian/Pacific Islander, NH	6.3	5.1	5.5	5.6	11.5	6.1	10.4	6.1
Other, NH	2.2	4.1	3.5	4.2	8.0	6.5	8.0	6.5
Mean Absolute Deviation (from IPEDS)	6.4		3.8		8.1		3.2	

^aCAS: Harvard College Alcohol Study;

^bIPEDS: Integrated Postsecondary Education System;

^cNCHRBBS: National College Health Risk Behavior Survey;

^dNCHA: National College Health Assessment;

^eNYAHS: National Young Adult Health Survey;

^fOnly the Harvard College Alcohol Study surveyed students under 18 years old – the other surveys excluded these students as part of their study designs;

^gIn the National College Health Assessment, multiple race/ethnicity responses were possible – cells total more than 100 percent