Social Psychiatry and Psychiatric Epidemiology (2024) 59:1919–1929
https://doi.org/10.1007/s00127-024-02704-3

RESEARCH

Discrepancies in prevalence estimates of suicidal ideation and attempts in 18- to 22-year-old US college students: a comparison of three surveys

James Aluri[1] · Daniel Eisenberg[2] · Mary T. Hoban[3] · Holly C. Wilcox[4] · Ramin Mojtabai[4]

Received: 30 August 2023 / Accepted: 14 June 2024 / Published online: 20 June 2024

Abstract

Purpose Suicidal thoughts and behaviors (STB) have been increasing among US college students. Accurate measurement of STB is key to understanding trends and guiding suicide prevention efforts. We aimed to compare the prevalence estimates of STB among college students from two campus-based surveys (the National College Health Assessment [NCHA] and the Healthy Minds Study [HMS]) and one general population study (the National Survey on Drug Use and Health [NSDUH]).

Methods Estimates were generated from the three surveys for past year suicidal ideation (PYSI) and past year suicide attempts (PYSA) among 18- to 22-year-old full-time college students. Data were combined from each survey to develop bivariate and multivariate regression models for odds of PYSI and PYSA.

Results Estimates for PYSI varied between the three surveys: 34.3% for NCHA, 15.0% for HMS, and 10.7% for NSDUH. Estimates for PYSA were 2.6% for NCHA, 1.6% for HMS, and 1.7% for NSDUH. After adjusting for demographic and educational characteristics, odds of PYSI remained significantly lower for HMS participants (aOR 0.31, 95% CI 0.29–0.33) and NSDUH participants (aOR 0.19, 95% CI 0.19–0.30) compared to NCHA participants. The odds of PYSA for HMS participants were lower than those for NCHA participants (aOR 0.63, 95% CI 0.54–0.73).

Conclusion Estimates of PYSI and PYSA vary between leading sources of data on college student mental health. The differences are likely related to question wording, survey implementation, as well as institutional and individual representation. Accounting for these differences when interpreting estimates of STB can help guide suicide prevention efforts.

Keywords Suicide, attempted · Suicidal ideation · College mental health · Higher education · Student health services · Health surveys

✉ James Aluri
james.aluri@jhmi.edu

Daniel Eisenberg
danieleisenberg@ucla.edu

Mary T. Hoban
mhoban@acha.org

Holly C. Wilcox
hwilcox1@jhmi.edu

Ramin Mojtabai
rmojtab1@jhu.edu

[1] Department of Psychiatry and Behavioral Sciences, Johns Hopkins University School of Medicine, Baltimore, MD, USA

[2] Department of Health Policy and Management, UCLA Fielding School of Public Health, Los Angeles, CA, USA

[3] Research Department, American College Health Association, Silver Spring, MD, USA

[4] Department of Mental Health, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA

Introduction

Suicide is the second leading cause of death for 15- to 24-year-olds in the United States [1]. While suicide deaths for college students are not distinguished, college students are 40% of 18- to 24-year-olds and represent a significant portion of this demographic [2]. The college years are critical developmentally, as students navigate new academic and social stressors away from their prior support system. These years are also an age during which mental disorders commonly manifest [3]. Suicidal thoughts and behaviors (STB) among college students had been increasing prior to the COVID-19 pandemic [4]. This trend was exacerbated by the pandemic due to disruptions in students' lifestyles and increased stress, leading to significant media coverage of recent student suicides [5–10].

Accurate measurement of STB among college students is key to understanding trends in student subgroups and factors such as social support and sexual orientation that are related to risk [11–13]. Identifying concerning trends and subgroups at higher risk allows colleges to target outreach efforts, suicide-screening programs [14], gatekeeper trainings [15], and government funded initiatives [16]. A meta-analysis of suicide prevention efforts on college campuses concluded that gatekeeper training programs improve suicide prevention knowledge in identifying students at risk for suicide and that interventions directed to at-risk groups reduce STB [17].

Information about STB in college students in the US primarily has come from two sources—campus-based surveys and general population surveys. Institutions can choose to participate in one of two major, national, campus-based mental health surveys—the Healthy Minds Study (HMS) and the National College Health Assessment (NCHA). HMS was launched in 2007 and assesses mental health symptoms, service utilization, and help-seeking behaviors [18]. The NCHA, managed by the American College Health Association, started in 2000 and is in its third iteration, the NCHA III [19]. Both rely on voluntary participation from students and institutions and are primarily used for internal evaluations. Institutions participating in these surveys receive individualized reports benchmarking their results among the pool of other participating colleges and universities. The characteristics (e.g., public versus private, enrollment numbers) and geographic locations of the institutions varies from year to year. The institutional sampling approach results in each year's survey representing thousands of students from the roughly one hundred institutions that participate and lacks students from the thousands of other institutions that did not participate.

Nationally representative population studies are the second main source of information about US college student mental health. These studies, funded by the US government, typically use probabilistic sampling to generate estimates for the entire US population. Annual general population studies that allow researchers to identify college students for sub-analyses include the National Survey of Drug Use and Health (NSDUH) and the Monitoring the Future Study (MTF) [20, 21]. MTF does not assess suicide, leaving NCHA, HMS, and NSDUH as the three primary sources for information on STB in college students [4, 22–24]. NSDUH samples participants by dwelling units (including college dorms) and not by institutional enrollment. Thus, NSDUH is likely to include students from a variety of housing contexts (i.e. dorms, on- or off-campus apartments, living with family). NSDUH student participants can also be drawn from geographically diverse institutions in both the public and private sector at different levels of post-secondary education (e.g., both community colleges and research universities).

Correspondence among these surveys' estimates of STB would increase confidence in the external validity of these estimates to represent STB in the US college-aged population. Conversely, discrepancies would require a closer examination of possible causes of the discrepancies. Clarifying the consistency of these estimates of STB and the possible sources of discrepancies will sharpen the interpretation of data used to direct suicide prevention efforts.

Prior work comparing prevalence of mental disorders in the general population between major surveys have led to refined estimates, improved interpretation, and recommendations for strengthening survey methods [25, 26]. A critical component of comparing prevalence estimates is adjusting for sample differences in age, sex, and race / ethnicity between participants of different surveys [26].

In this study, we aimed to compare the prevalence estimates of STB among college students from two campus-based surveys and one general population study. Specifically, we sought:

1. To compare the prevalence of past year suicidal ideation (PYSI) and past year suicide attempt (PYSA) among full-time, 18- to 22-year-old college students who participated in NCHA, HMS, and NSDUH in 2021.
2. To evaluate whether any discrepancies in PYSI and PYSA among full-time, 18- to 22- year-old college students who participated in NCHA, HMS, and NSDUH in 2021 persist after adjusting for differences in the population composition of these surveys with regard to age, sex, sexual orientation, race, and educational level.

Methods

Study design

The National College Health Assessment III (NCHA III) was developed by the American College Health Association (ACHA) and administered annually starting in Fall 2019. It is an annual survey conducted at postsecondary institutions (colleges and universities) that volunteer to participate. NCHA invites a random sample from each institution, except at smaller institutions where all students are invited to participate. NCHA does not employ weights, strata, or cluster in its survey design elements to adjust for non-response biases or other factors. The survey is confidentially self-administered online. Data files are then anonymized before analysis.

The Healthy Minds Study (HMS) is another annual survey implemented by self-selecting, postsecondary institutions and uses the same approach as the NCHA by random sampling students at institutions, except for smaller institutions where all are invited. To compensate for non-response bias, non-response weights are created to account for over-representation of certain characteristics among respondents. Prior to Fall 2021, weights took into account gender, race / ethnicity, academic level, and grade point average [27]. Starting Fall 2021, female versus male sex was the only characteristic accounted for in weighting as sex was consistently the strongest predictor of response [23]. As with the NCHA III, the survey is self-administered anonymously online.

The National Survey on Drug Use and Health (NSDUH) is an annual, government-sponsored survey of substance use and mental health in the US used to provide nationally representative data. NSDUH utilizes a probabilistic, multi-stage sampling strategy as described in its annual methodological summary [28]. The methodological summary also details how weights are calculated. In brief, weights combine both the probability of selection due to the survey sampling design and adjustment factors for non-response, extreme weights, and post-stratification to make the sample representative of the US 2020 Census population. NSDUH had two modes of administration in 2021—self-administration online and administration with support from a field interviewer. The field interviewing process allows participants to self-administer portions of the exam related to sensitive topics, including suicide [29]. The proportion of participants using the online version declined by quarter from 76.6% in Quarter 1 to 41.5% in Quarter 4.

Our study was deemed to not constitute human subjects research by the Johns Hopkins School of Public Health Institutional Review Board (application ID IRB00024452).

Sample

For each study, we restricted the sample to students who were 18- to 22-years-old, enrolled full-time in college, and participated in the respective surveys in 2021. For NCHA and HMS, whose data are aggregated by academic year, we combined two academic years of data (2020–2021 and 2021–2022) then selected participants from Spring 2021 and Fall 2021. HMS response rates were 15% for Spring 2021 and 18% for Fall 2021 [27]. Mean response proportions for NCHA participating institutions were 13% for both Spring 2021 and Fall 2021 [30, 31]. The NSDUH weighted interview response rate was 46.2% [28]. Final unweighted sample sizes were 76,618 students for the NCHA; 77,424 for HMS; and 2,910 for NSDUH. NSDUH is a household survey and achieves reasonable coverage of college students. Its weighted 2021 sample estimates there were 20.3 million college students in the US, including 10.8 million female students and 7.3 million part-time students. Data from the Integrated Postsecondary Education Data System shows that there were 18.7 million college students in the US, including 10.9 million female students and 7.2 million part-time students [32].

Variables

Dependent variables

Suicidal ideation In NCHA III, participants were asked, "How often have you thought about killing yourself in the past year?" Responses included "Never", "Rarely (1 time)", "Sometimes (2 times)", "Often (3–4 times)", "Very often (5 or more times)." We re-coded "Never" to "No", and all other answers to "Yes." In HMS, participants were asked, "In the past year, did you ever seriously think about attempting suicide?" Similarly, in NSDUH, participants were asked, "At any time in the past 12 months, that is from [the date for one year ago] up to and including today, did you seriously think about trying to kill yourself?" Both in HMS and NSDUH participants could answer "Yes" or "No" to these questions.

Suicide attempts In NCHA III, all participants were asked, "Within the last 12 months, have you attempted suicide?" HMS used a survey skip logic and only participants who reported having suicidal ideation in the past year were asked, "In the past year, did you attempt suicide?" In NSDUH, all participants were asked, "During the past 12 months, did

you try to kill yourself?" For all three surveys, the participants could answer "Yes" or "No" to these questions.

Independent variables

Age, sex, sexual orientation, race / ethnicity, year in college, and housing data were obtained from each survey. For a complete description of how these variables were obtained and coded by each survey, please see Supplemental Table 1. Psychiatric distress over the past 30 days (as measured by the Kessler 6, a validated six-question scale that has been described elsewhere [33]) was assessed in NCHA and NSDUH.

Weights, strata, cluster

NCHA does not employ weights, strata or cluster in its survey design. HMS provides weights in the public data file. NSDUH's public use files include weights, strata, and cluster.

Recoding for combined analysis

Age and education level were recoded to match NSDUH's variables given that NSDUH had the least granular data. The HMS 2022 study did not differentiate non-answers from negative answers when coding many demographic variables that were asked about in a "select all that apply" format. For those questions, our analysis considered all non-affirmative answers as negative answers. Variables that were asked about as "select all that apply" in some surveys, but not others (i.e. race variables from NCHA and HMS, and sexual orientation variables from HMS) were recoded as a mutually exclusive categorical variable to match NSDUH's recoded race and sexual orientation variables.

Unique institutional IDs were used as non-overlapping "clusters" for the NCHA and HMS. Original NSDUH strata were used. To facilitate combining data from the surveys for the analyses, HMS and NCHA data were each assigned a unique stratum and NCHA participants were given a weight of 1. HMS and NSDUH weights were adjusted to a mean of 1 for each survey. All participant data were combined into a single dataset using the standardized variables and a 'Survey source' variable was added to indicate which survey the participant data was obtained from.

Analysis

Descriptive statistics

Descriptive statistics were generated for the participants of each survey. Unweighted counts were included for each survey. Proportions for NCHA were not weighted. HMS data were described using provided weights, consistent with prior analyses of these data [23]. NSDUH data were described using provided weight, cluster, and strata information.

Prevalence estimates

Prevalence estimates for PYSI and PYSA were generated using relevant survey design information for each survey (no survey elements for NCHA; weights only for HMS; weights, cluster, and strata for NSDUH). For NCHA, 95% Confidence Intervals (CIs) were calculated with R commands that relied on the Wilson Score method [34]. For HMS and NSDUH, which used complex survey design elements, CIs were calculated by fitting a logit regression, computing a Wald-type interval on the log-odds scale, and then transforming the interval back to the probability scale. HMS assumes that those who attempted suicide must have had suicidal ideation and does not ask those who did not have PYSI about PYSA. Thus, HMS's weighted prevalence estimates for PYSA exclude those who would have reported attempts without reporting ideation.

Logistic regression models

Bivariate logistic regression models were built for the outcomes of PYSI and PYSA using a list of explanatory variables that included variables that could be standardized across the three surveys, i.e. age, sex, sexual orientation, race / ethnicity, and year in college. We also included the 'source' variable, which indicated which survey the data was from. Multivariate models were then built including all covariates to adjust for differences between participants in each survey.

The regression models included the survey weights, strata, and cluster as described in Section "Recoding for combined analysis". Odds ratios were generated by exponentiating the coefficients from the models. For all analyses, 95% confidence intervals were generated using Wald-type confidence intervals.

Given the lack of commensurability of some variables across the three surveys, we excluded certain low prevalence categories from the dataset used for the regression analyses. Specifically, when standardizing the dataset for the regression models we removed participants from the following categories: sex reported as intersex ($n = 149$); sexual

orientation reported as asexual, pansexual, queer, questioning, or other ($n = 17,546$); race reported as Middle Eastern, North African, Arab or 'Other' ($n = 2,588$); educational level reported as other and non-degree seeking ($n = 1,424$).

Results

Sample characteristics

Throughout the results section, all HMS and NSDUH percentages are weighted, but NCHA percentages are not.

The study samples from each of the three survey sources appeared to differ in terms of sex, sexual orientation, race, educational level and housing (Table 1). NSDUH had the lowest proportion of female respondents (52.5%) compared to 69.3% in NCHA and 59.4% in HMS. NSDUH had a higher percentage of participants who identified as heterosexual or straight (81.7%) compared to 74.3% for NCHA and 75.2% for HMS. NCHA had the highest proportion of Asian students (18.4%) compared to 10.6% for HMS and 8.5% for NSDUH. NCHA also had the lowest proportion of Black or African American students (4.9%) compared to 14.0% for HMS and 12.8% for NSDUH. HMS had the highest percentage of white students (68.9%) compared to 63.1% for NCHA and 56.3% for NSDUH. NSDUH had the highest percentage of Hispanic students (19.8%) compared to 18.1% for NCHA and 12.2% for HMS. After standardization of the data set, the survey samples differed in sex, race, and year of education—but not sexual orientation (Supplemental Table 2). HMS's estimate of 59.4% of students being female is closest to the 58.4% female students provided by the National Center for Education Statistics in Fall 2021 [32]. NSDUH's distribution of the major racial and ethnic groups (White = 56.3%, Hispanic or Latinx = 19.8%, Black or African American = 12.8%, and Asian or Asian American = 8.5%) closely matched actual data from Fall 2021 enrollment (53.4%, 20.6%, 13.1%, and 7.6%, respectively).

The distribution among types of housing appeared similar between HMS and NCHA. NSDUH's housing information was not commensurate with the other two surveys but appeared to have a much lower percentage of students in dormitories (1.1%). Around 40% of participants in the other two surveys indicated they lived on campus or in university housing, which was the closest category to NSDUH's dormitory variable.

Prevalence estimates of PYSI and PYSA

NCHA had the highest prevalence estimate of PYSI (34.3%) compared to 15.0% for HMS and 10.7% for NSDUH (Table 2). NCHA still had the highest prevalence of PYSI (20.8%) even if students who only thought about killing themselves once in the past year (13.4%) were excluded. NCHA also had the highest estimate of PYSA at 2.6% compared to 1.6% for HMS and 1.7% for NCHA. The HMS and NSDUH estimates were closer to each other than to NCHA for both PYSI and PYSA (Fig. 1).

Multivariable comparison of PYSI and PYSA by survey type

After adjusting for demographic and educational characteristics of the samples, discrepancies in estimates of PYSI and PYSA remained (Table 3). The odds of reporting PYSI were significantly lower for HMS participants (aOR 0.33, 95% CI 0.30–0.35) and NSDUH participants (aOR 0.25, 95% CI 0.20–0.32) compared to NCHA participants. The difference in adjusted odds of PYSI between HMS and NSDUH participants was also significant at the level of $p < 0.05$. The adjusted odds ratios for PYSA compared to NCHA participants (the reference group) were closer to 1 for both HMS (aOR 0.66, 95% CI 0.56–0.77) and NSDUH participants (aOR 0.67, 0.40–1.12) than the adjusted odds ratios for PYSI from the same surveys (aOR 0.33 and 0.25, respectively). Interpreting the 95% CIs for the aOR of PYSA, the difference in odds of PYSA was significant for HMS, but not for NSDUH, compared to NCHA. In a sensitivity analysis that applied HMS's skip logic to the NCHA and NSDUH populations, both odds ratios moved even closer to 1 (Supplemental Table 3), however, the HMS odds remained significantly different (aOR 0.83, 95% CI 0.72–0.95).

Several other characteristics were also significantly associated with PYSI and PYSA (Table 3). Male sex was associated with significantly lower odds of PYSI compared to female sex (aOR 0.91, 95% CI 0.87–0.95), but significantly higher odds of PYSA (aOR 1.25, 95% CI 1.10–1.41). Students identifying as Gay, Lesbian, or Bisexual had significantly higher adjusted odds of both PYSI and PYSA compared to students identifying as heterosexual.

In terms of race, higher odds of PYSI were found for students identifying as Black (aOR 1.18, 95% CI 1.09–1.29), Native Hawaiian or Pacific Islander (aOR 1.87, 95% CI 1.30–2.69), Asian (aOR 1.11, 95% CI 1.04–1.19), and Multi-racial (1.30, 95% CI 1.22–1.39). Students identifying as Black, Asian, Multi-racial, and Hispanic had significantly higher odds of PYSA. Graduate students or undergraduate students in their fourth or higher year of studies had lower odds of PYSI (aOR 0.88, 95% CI 0.83–0.94) and PYSA (aOR 0.77, 95% CI 0.63–0.94) compared to first-year undergraduate students. The relationship between the demographic and educational variables and suicide outcomes were similar across the three surveys (Supplemental Table 4).

Table 1 Sociodemographic and educational characteristics among 18- to 22-year-old, full-time enrolled college students who participated in NCHA, HMS, and NSDUH

	NCHA $n=76,618$			HMS $n=77,424$			NSDUH $n=2,910$	
	n	%	Category Σ	*n*	Weighted %	Category Σ	*n*	Weighted %
Age								
18	13,735	17.9	65.4[a]	14,117	18.7	65.6[a]	1,826	66.1[a]
19	18,682	24.4		18,394	23.8			
20	17,703	23.1		18,429	23.1			
21	16,390	21.4	34.6[a]	16,429	21.4	34.4[a]	1,084	33.9[a]
22	10,108	13.2		10,055	13.0			
Sex								
Female	53,085	69.3		56,133	59.4		1,233	52.5
Male	23,465	30.6		21,202	40.6		1,677	47.5
Intersex	15	0.0		27	0.0			-
Sexual orientation								
Heterosexual / Straight	56,746	74.3		56,983	75.2		2,324	81.7
Gay / Lesbian	3,074	4.0		3,478	5.0		116	4.0
Bisexual	9,421	12.3		11,102	13.7		415	14.3
Asexual, pansexual, queer, questioning, other	7,117	9.3		8,848	11.0			-
Race / Ethnicity								
American Indian or Native Alaskan	1,337	1.7		997	1.5		13	0.1
Asian or Asian American	14,089	18.4		9,863	10.6		284	8.5
Black or African American	3,746	4.9		8,783	14.0		276	12.8
Hispanic or Latinx	13,837	18.1		8,578	12.2		585	19.8
Middle Eastern, North African, or Arab	1,315	1.7		1,588	1.8		-	-
Native Hawaiian or Other Pacific Islander	588	0.8		462	0.6		6	0.1
White	48,359	63.1		54,639	68.9		1,600	56.3
Multiracial	4,206	5.5			-		146	2.5
Other	701	0.9		799	1.2		-	-
Educational level								
Undergraduate: 1st year	22,126	28.9		21,567	29.7		771	26.9
Undergraduate: 2nd year	18,103	23.6	48.0[a]	20,087	26.2	48.5[a]	1,401	49.5[a]
Undergraduate: 3rd year	18,677	24.4		18,326	22.3			
Undergraduate: 4th year	13,858	18.1	22.8[a]	12,948	16.2	19.5[a]	738	23.6[a]
Undergraduate: 5th year or more	1,321	1.7		787	0.9			
Graduate	2,305	3.0		2,174	2.4			
Other / non-degree seeking	207	0.3		1,119	2.3			-
Housing								
Campus or university housing	28,903	38.3		28,639	42.7			-
College dormitory	-	-			-		36	1.1
Parent / guardian / other family member's home	23,224	30.8		22,711	30.9			-
Off campus or other non-university housing	22,816	20.2		23,489	23.3			-
Other (e.g., unstable housing, homeless)	530	0.7		634	1.1			-
Fraternity or sorority	-	-		792	0.8			-
Other cooperative housing	-	-		932	1.2			-

Shown are the sociodemographic and educational characteristics for each of the survey samples. All *n* values are unweighted counts. NSDUH's variables were less granular than the NCHA and HMS variables and table rows were merged in this column to match combinations of categories from the other survey samples. If one survey did not capture a particular variable it was signified with a dash

[a]These numbers represent proportions for combined categories indicated by contiguous, same-shaded rows

In a sensitivity analysis, we evaluated for difference in psychiatric distress using a standardized measure shared by NCHA and NSDUH (the Kessler 6 or K6). The mean distress score was significantly higher for NCHA participants (9.11, 95% CI 9.07–9.15) than for NSDUH participants (7.14, 95% CI 6.71–7.57) (Supplemental Table 5). Adjusting for K6 scores in a regression model for PYSI and PYSA did not eliminate the differences between surveys, as NSDUH participants still had lower odds of reporting PYSI (aOR 0.22, 95% CI 0.17–0.27) and PYSA (aOR 0.56, 95%

Table 2 Estimates of Prevalence of past year Suicidal Ideation and past year Suicide Attempts among 18- to 22-year-old, full-time enrolled college students who participated in NCHA, HMS, and NSDUH

	NCHA		HMS		NSDUH	
	%	(95% CI)	Weighted %	(95% CI)	Weighted %	(95% CI)
Suicidal ideation	34.3	(34.0–34.7)	15.0	(14.6–15.4)	10.7	(8.8–13.0)
Suicide attempts	2.6	(2.5–2.7)	1.6	(1.5–1.8)	1.7	(1.0–2.8)

Shown are the estimates of the prevalence of suicidal ideation and suicide attempts by survey and their 95% confidence intervals. These figures represent the full datasets for each survey, prior to the data standardization that required removal of certain categories of incommensurate data

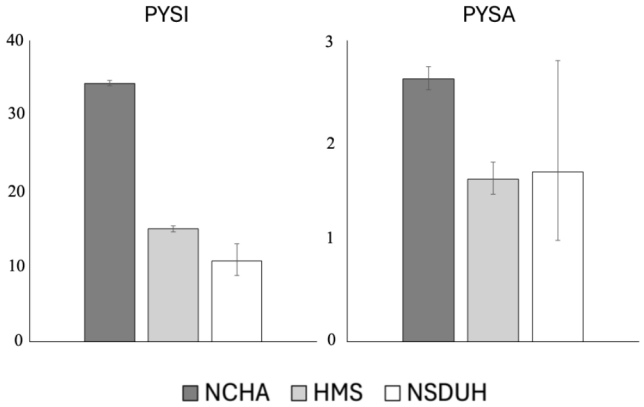

Fig. 1 Prevalence estimates, by survey, for past year suicidal ideation (PYSI) and past year suicide attempts (PYSA)

CI 0.33–0.94) adjusting for all other covariates (Supplemental Table 6).

Discussion

This study is the first to compare prevalence estimates of STB among college students between campus-based and general population surveys. The estimates of PYSI varied significantly between surveys, even after adjusting for underlying differences in age, sex, sexual orientation, race/ethnicity, and year in college. For PYSA, the HMS and NSDUH estimates were lower than the NCHA estimate, though only the NCHA-HMS difference was significant.

The potential reasons for these discrepancies can be divided into differences in survey questions, implementation, and representation. The NCHA question for PYSI is qualitatively different than the other two surveys. Using a question in the validated Suicide Behavior Questionnaire-Revised scale (SBQ-R), the NCHA asks about frequency of suicidal ideation ("How often…", while the other surveys ask "Did you…") and does not use the word "seriously," a word the other two surveys include. This question could include intrusive thoughts about suicide—a different phenomenological experience than intending to kill oneself [35]. The absence of the word "seriously" in the NCHA question represents a 'broad' construct of suicide and likely yields higher PYSI estimates than those generated by the other surveys' 'narrow' constructs [36, 37].

With respect to survey implementation, HMS's lower estimate for PYSA is likely due, at least in part, to the question about SA being only asked of students who report PYSI. In the other surveys, students reported PYSA without having PYSI, and our sensitivity analysis applying this logic to the other surveys mitigated the magnitude of the differences in odds of reporting PYSA (though the HMS versus NCHA comparison remained significant). Another implementation factor might be mode of administration. Some NSDUH participants self-administer the survey parallel to how the NCHA and HMS surveys are conducted. However, other NSDUH participants take the survey in person. While NSDUH allows in-person participants to self-administer sensitive parts of the exam, conducting assessments in the home with an interviewer and family members present might result in underreporting of stigmatized behaviors [28]. For example, NSDUH participants might fear that reporting STB would precipitate a crisis intervention.

Some variability might be explained by differences in how well the surveys represent the college student population. These differences are expected as the surveys have different purposes and sampling strategies. First, the voluntary nature of institutional enrollment in the campus-based surveys might result in unequal representation of types of US postsecondary institutions between the three surveys. This might be particularly important given that students attending public, less competitive, and less residential schools have higher odds of reporting suicidal ideation compared to students at private, more competitive, and more residential schools [38].

Second, NCHA does not weight its responses and had a higher prevalence of female participants, which was associated with PYSI. However, the lack of weighting cannot fully explain the discrepancy as our regression models still showed a significant difference in odds of suicidal ideation after adjusting for sex.

Third, non-response biases might influence the results. NSDUH had the lowest estimate of PYSI and the highest response rates. A meta-analysis of STB among college students (median age 21.4 years) found an association between lower response rates and higher prevalence estimates of suicidal ideation among surveys that used a broad definition of suicidal ideation [37]. This is likely because those with suicidal ideation are overrepresented in self-selecting surveys

Table 3 Unadjusted and adjusted odds ratios for past year suicidal ideation and past year suicide attempts by demographic and educational characteristics among pooled survey participants

	Past year suicidal ideation		Past year suicide attempts	
	OR (95% CI)	aOR (95% CI)	OR (95% CI)	aOR (95% CI)
Age category				
18–20	REF	REF	REF	REF
21–22	0.88 (0.84, 0.92)	0.96 (0.91, 1.01)	0.74 (0.67, 0.83)	0.87 (0.75, 1.01)
Sex				
Female	REF	REF	REF	REF
Male	0.72 (0.69, 0.77)	0.91 (0.87, 0.95)	1.07 (0.95, 1.21)	1.25 (1.10, 1.41)
Sexual orientation				
Heterosexual / Straight	REF	REF	REF	REF
Gay, lesbian	2.44 (2.25, 2.67)	2.66 (2.46, 2.87)	2.58 (2.09, 3.20)	2.56 (2.06, 3.17)
Bisexual	3.09 (2.94, 3.30)	3.24 (3.09, 3.40)	2.49 (2.22, 2.79)	2.62 (2.34, 2.94)
Race / Ethnicity				
White, non-Hispanic	REF	REF	REF	REF
Black, non-Hispanic	0.86 (0.77, 0.97)	1.18 (1.09, 1.29)	1.49 (1.21, 1.83)	1.73 (1.41, 2.11)
AI or NA, non-Hispanic	1.27 (0.92, 1.74)	1.01 (0.76, 1.35)	1.41 (0.65, 3.07)	1.26 (0.59, 2.69)
NH or PI, non-Hispanic	1.92 (1.37, 2.68)	1.87 (1.30, 2.69)	1.65 (0.69, 3.99)	1.68 (0.69, 4.08)
Asian, non-Hispanic	1.24 (1.13, 1.36)	1.11 (1.04, 1.19)	1.23 (1.06, 1.43)	1.21 (1.05, 1.40)
Multi-racial, non-Hispanic	1.53 (1.42, 1.65)	1.30 (1.22, 1.39)	1.39 (1.14, 1.68)	1.24 (1.03, 1.51)
Hispanic	1.11 (1.02, 1.20)	0.96 (0.90, 1.01)	1.29 (1.12, 1.48)	1.20 (1.05, 1.38)
Year				
First-year undergrad	REF	REF	REF	REF
Second- or third-year undergrad	0.99 (0.95, 1.04)	1.01 (0.97, 1.06)	0.87 (0.77, 0.99)	0.92 (0.81, 1.04)
Fourth-year undergrad or more (including grad students)	0.88 (0.83, 0.94)	0.88 (0.83, 0.94)	0.67 (0.58, 0.77)	0.77 (0.63, 0.94)
Survey				
NCHA	REF	REF	REF	REF
HMS	0.33 (0.30, 0.35)	0.31 (0.29, 0.33)	0.66 (0.56, 0.77)	0.63 (0.54, 0.73)
NSDUH	0.25 (0.20, 0.32)[b]	0.24 (0.19, 0.30)[b]	0.67 (0.40, 1.12)[c]	0.63 (0.38, 1.05)[c]

Shown are the unadjusted odds ratios (ORs) and adjusted odds ratios (aORs) for suicidal ideation and suicide attempts by demographic and educational characteristics among a pool of participants from all three studies

[b]Comparison of NSDUH and HMS has p value < 0.05

[c]Comparison of NSDUH and HMS has p value > 0.05

with small response rates. In our data, we found that average distress was higher in NCHA compared to NSDUH, as measured by Kessler 6 scores. However, this does not fully explain the discrepancies as the sensitivity analysis adjusting for K6 scores still found significant differences between NCHA and NSDUH estimates for PYSI and PYSA. While this overrepresentation of students with suicidal ideation might threaten the results' external validity to the broader college population, these students are precisely the people whose data will be most helpful to guide suicide prevention efforts.

NSDUH's estimate for suicidal ideation was closest to the PYSI estimate of 10.6% from the meta-analysis of STB among college students [37]. While the meta-analysis included non-US studies, US participants comprised 88.5% of the pooled sample and 89.3% of US participants were from NCHA, HMS, or NSDUH. These prevalence estimates were likely lower than the NCHA and HMS estimates because (1) narrower measures of suicidal ideation were used in prior versions of the NCHA and (2) suicidal ideation has been increasing over time. The meta-analysis included data and studies from 2000 to 2016 for the NCHA and from 2007 to 2016 for the HMS [39].

We acknowledge several limitations to our study. First, the social and demographic variables were not directly comparable across surveys which led to the exclusion of certain groups from the study (students identifying as intersex, asexual, pansexual, queer, questioning, or Middle Eastern, north African, Arab or Other Race / Ethnicity) which could limit our ability to compare assessments of STB among these surveys for those groups. This also limited our ability to identify whether there were differences in socioeconomic characteristics, such as family income, and how they might impact the variation in estimates. Second, to compare a similar profile of college students across studies, we limited our analysis to 18- to 22-year-olds who were full-time enrolled to eliminate an extra source of variability (heterogeneity among part-time students). Third, we were unable to adjust for housing given that NSDUH's housing variables were not commensurate with the other surveys. However, we do not

believe that NSDUH's smaller proportion of students from college dorms explains the lower estimate of PYSI, as prior studies have not found a positive correlation between living in a dorm and suicidal ideation [12, 40].

The strengths of our study include robust statistical methods to adjust for differences in demographic and social characteristics and a novel combination of multiple datasets to better understand relationships between individual characteristics and STB. These methods could be applied to investigate other discrepancies between surveys, such as non-suicidal self-injurious behavior (NSSI). HMS estimated NSSI prevalence at 28% in the 2021–2022 academic year, while NCHA estimated it was 9.6% for Fall 2021 [23, 31].

We outline four implications of our study for institutional leaders and researchers. First, the absence of consistent estimates of STB does not necessarily lead to the conclusion that the estimates are internally invalid. Each survey could have valid estimates for their respective construct of STB among the participants at the institutions that were assessed. Nor do our findings preclude the use of these surveys for assessing STB for institutional benchmarking and priority-setting. Encouragingly, relationships between demographic variables and STB were similar among the surveys suggesting that each survey can identify groups of concern and inform targeted suicide prevention efforts. However, we urge against interpreting an estimate from any one survey as representing a single, externally-valid construct of STB for US college students. Second, researchers should interpret NCHA's broad construct of PYSI cautiously given that it lacks the word "seriously" (which implies clinically-relevant intent) and recognize the likely underestimation of PYSA from HMS due to its skip logic. NCHA's research team should consider whether an alternative measure for PYSI more closely aligns with its goals. Starting with its 2023–2024 survey, HMS is removing its skip logic for assessing suicide plans and attempts based on the findings from the present study. Third, standardization of measures for STB and other mental illness might help disentangle the role that different constructs play in these discrepant estimates. For example, of these three surveys, only HMS includes the PHQ-9, which is endorsed by the Common Measures in Mental Health Science Initiative [41]. Fourth, each survey's research team should consider whether external validity to the broader population of US college students aligns with the survey's goals and what it would take to improve that validity. For NCHA and HMS, this would entail efforts to broaden institutional and individual participation. The NCHA research team is actively exploring the possibility of adding non-response weighting adjustments, similar to those of HMS, to future versions of the survey. NSDUH's probabilistic, multi-stage sampling that does not rely on institutional enrollment and its higher participation rate are advantages. However, its undersampling from college dorms is a limitation and the influence of other aspects of its sampling design on the generalizability to the broader college population needs to be evaluated.

Conclusion

Estimates of suicidal ideation and suicide attempts among 18- to 22-year-old, full-time college students vary significantly among campus-based and general population surveys even after adjusting for several sociodemographic characteristics. The discrepancies appear to stem from differences in question wording, survey implementation, as well as institutional and individual representation. Understanding these differences can improve the measurement and interpretation of STB estimates, which inform institutional benchmarking, priority setting, and designing suicide prevention initiatives.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00127-024-02704-3.

Acknowledgements Not applicable.

Author contributions James Aluri: conceptualization, methodology, software, validation, formal analysis, data curation, writing-original draft, review & editing, visualization, project administration. Daniel Eisenberg: conceptualization, methodology, resources, writing – review & editing, supervision. Mary T. Hoban: conceptualization, methodology, resources, writing – review & editing, supervision. Holly C. Wilcox: methodology, writing – review & editing, supervision. Ramin Mojtabai: conceptualization, methodology, writing – original draft, review & editing, supervision.

Funding Dr. Aluri and Dr. Wilcox receive research support from the American Foundation for Suicide Prevention. Dr. Wilcox is a member of the Research Grants Committee and Scientific Advisor at the AFSP. Dr. Eisenberg co-founded the Healthy Minds Study and is a Principal Investigator of the Healthy Minds Study. Dr. Hoban is employed by the American College Health Association and directs the National College Health Assessment. Dr. Mojtabai has received royalties from UpToDate and Medscape and has provide expert consultation regarding social media litigation on behalf of the plaintiffs. The authors did not receive support from any organization for the submitted work.

Data availability Data from NCHA and HMS are available after approval of appropriate data request procedures. NSDUH public use files are publicly available online.

Declarations

Ethical approval The manuscript does not contain clinical studies or patient data. Our study was deemed not to be human subjects research by the Johns Hopkins School of Public Health Institutional Review Board (application ID IRB00024452).

References

1. National Center for Health Statistics (2021) Deaths: leading causes for 2019 - supplemental tables. National Vital Statistics Reports: National Center for Health Statistics
2. Statistics NCfE (2023) College Enrollment Rates. Condition of Education. U.S. Department of Education, Institute of Education Sciences
3. Pedrelli P, Nyer M, Yeung A, Zulauf C, Wilens T (2015) College students: mental health problems and treatment considerations. Acad Psychiatry 39:503–511
4. National Academies of Sciences Engineering and Medicine (2021) Mental Health, Substance Use, and Wellbeing in Higher Education: supporting the whole student. The National Academies, Washington, DC
5. Conrad RC, Koire A, Pinder-Amaker S, Liu CH (2021) College student mental health risks during the COVID-19 pandemic: implications of campus relocation. J Psychiatr Res 136:117–126
6. Park JH, Bui K (2022) Mental health of undergraduates one year after the start of the COVID-19 pandemic: findings from the national college health assessment III. J Am Coll Health. 1–4
7. Thompson MP, Tyson JS, Hege A, Seitz C (2023) COVID-related stress, risk for suicidal behavior, and protective factors in a national sample of college students. J Am Coll Health. 1–8
8. Hartocollis A (2021) Colleges Fear Mental Health Crisis Amid Covid Surge. The New York Times
9. Seidman L (2022) After soccer star Katie Meyer's death, Stanford vows to bolster mental health services. Los Angeles Times
10. Shammas B UNC cancels classes after reported suicide and attempt, citing 'mental health crisis'. The Washington Post2021
11. Wilcox HC, Arria AM, Caldeira KM, Vincent KB, Pinchevsky GM, O'Grady KE (2010) Prevalence and predictors of persistent suicide ideation, plans, and attempts during college. J Affect Disord 127(1):287–294. https://doi.org/10.1016/j.jad.2010.04.017
12. Mortier P, Auerbach RP, Alonso J, Bantjes J, Benjet C, Cuijpers P et al (2018) Suicidal thoughts and behaviors among first-year college students: results from the WMH-ICS project. J Am Acad Child Adolesc Psychiatry 57(4):263–273 e1
13. Li W, Dorstyn DS, Jarmon E (2020) Identifying suicide risk among college students: a systematic review. Death Stud 44(7):450–458
14. Moffitt LB, Garcia-Williams A, Berg JP, Calderon ME, Haas AP, Kaslow NJ (2014) Reaching graduate students at risk for suicidal behavior through the interactive screening program. J Coll Student Psychother 28(1):23–34
15. Lipson SK, Speer N, Brunwasser S, Hahn E, Eisenberg D (2014) Gatekeeper training and access to mental health care at universities and colleges. J Adolesc Health 55(5):612–619
16. Goldston DB, Walrath CM, McKeon R, Puddy RW, Lubell KM, Potter LB et al (2010) The Garrett Lee Smith Memorial Suicide Prevention Program. Suicide Life-Threatening Behav 40(3):245–256. https://doi.org/10.1521/suli.2010.40.3.245
17. Wolitzky-Taylor K, LeBeau RT, Perez M, Gong-Guy E, Fong T (2020) Suicide prevention on college campuses: what works and what are the existing gaps? A systematic review and meta-analysis. J Am Coll Health 68(4):419–429. https://doi.org/10.1080/07448481.2019.1577861
18. Healthy Minds Network The healthy minds study - student survey. https://healthymindsnetwork.org/hms/ Accessed
19. American College Health Association About NCHA: Survey. https://www.acha.org/NCHA/About_ACHA_NCHA/Survey/NCHA/About/Survey.aspx?hkey=7e9f6752-2b47-4671-8ce7-ba7a529c9934 Accessed
20. Schulenberg JE, Patrick ME, Johnston LD, O'Malley PM, Bachman JG, Miech RA (2021) Monitoring the Future National Survey Results on Drug Use, 1975–2020. Volume II, College Students & Adults Ages 19–60. Institute for social research
21. Center for Behavioral Health Statistics and Quality (2021) National Survey on Drug Use and Health Public Use File Codebook. Rockville, MD: Substance Abuse and Mental Health Services Administration; 2022
22. Han B, Compton WM, Eisenberg D, Milazzo-Sayre L, McKeon R, Hughes A (2016) Prevalence and mental health treatment of suicidal ideation and behavior among college students aged 18–25 years and their non-college-attending peers in the United States. J Clin Psychiatry 77(6):20285
23. Eisenberg D, Lipson SK, Heinze J, Zhou S (2023) The Healthy Minds Study 2021–2022 Data Report. Ann Arbor, Michigan: Healthy Minds Network for Research on Adolescent and Young Adult Mental Health
24. ACHA. UNDERGRADUATE STUDENT Reference Group (2023) Data report fall 2022. American College Health Association
25. Narrow WE, Rae DS, Robins LN, Regier DA (2002) Revised prevalence estimates of mental disorders in the United States: using a clinical significance criterion to reconcile 2 surveys' estimates. Arch Gen Psychiatry 59(2):115–123
26. Regier DA, Kaelber CT, Rae DS, Farmer ME, Knauper B, Kessler RC et al (1998) Limitations of diagnostic criteria and assessment instruments for mental disorders: implications for research and policy. Arch Gen Psychiatry 55(2):109–115
27. Eisenberg D, Lipson SK, Heinze J, Zhou S (2021) The Healthy Minds Study 2021 Winter/Spring Data Report. Ann Arbor, Michigan: Healthy Minds Network for Research on Adolescent and Young Adult Mental Health
28. Center for Behavioral Health Statistics and Quality (2021) National Survey on Drug Use and Health (NSDUH): Methodological summary and definitions. Substance Abuse and Mental Health Services Administration; 2022
29. Statistics CfBH Q (2021) National Survey on Drug Use and Health (NSDUH): Final CAI Specifications for Programming. Substance Abuse and Mental Health Services Administration; 2020
30. American College Health Association (2021) Spring 2021: Reference Group Data Report. American College Health Association
31. American College Health Association (2021) Fall 2021: Reference Group Data Report. American College Health Association
32. (IPEDS) IPEDS (2022) Table 303.10. Total fall enrollment in degree-granting postsecondary institutions, by attendance status, sex of student, and control of institution. Selected years, 1947 through 2031. Fall Enrollment Survey: U.S. Department of Education, National Center for Education Statistics
33. Kessler RC, Andrews G, Colpe LJ, Hiripi E, Mroczek DK, Normand S-L et al (2002) Short screening scales to monitor population prevalences and trends in non-specific psychological distress. Psychol Med 32(6):959–976
34. Newcombe RG (1998) Two-sided confidence intervals for the single proportion: comparison of seven methods. Stat Med 17(8):857–872
35. Osman A, Bagge CL, Gutierrez PM, Konick LC, Kopper BA, Barrios FX (2001) The suicidal behaviors Questionnaire-revised (SBQ-R): validation with clinical and nonclinical samples. Assessment 8(4):443–454
36. Ammerman BA, Burke TA, Jacobucci R, McClure K (2021) How we ask matters: the impact of question wording in single-item measurement of suicidal thoughts and behaviors. Prev Med 152:106472
37. Mortier P, Cuijpers P, Kiekens G, Auerbach R, Demyttenaere K, Green J et al (2018) The prevalence of suicidal thoughts and behaviours among college students: a meta-analysis. Psychol Med 48(4):554–565

38. Lipson SK, Gaddis SM, Heinze J, Beck K, Eisenberg D (2015) Variations in Student Mental Health and Treatment Utilization across US Colleges and universities. J Am Coll Health 63(6):388–396. https://doi.org/10.1080/07448481.2015.1040411
39. Mortier P, Cuijpers P, Kiekens G, Auerbach R, Demyttenaere K, Green J et al (2018) The prevalence of suicidal thoughts and behaviours among college students: a meta-analysis—supplementary materials 3. Psychol Med 48(4):554–565
40. Eisenberg D, Gollust SE, Golberstein E, Hefner JL (2007) Prevalence and correlates of depression, anxiety, and suicidality among university students. Am J Orthopsychiatry 77(4):534–542
41. Farber GK, Gage S, Kemmer D (2023) A collaborative effort to establish Common Metrics for Use in Mental Health. JAMA Psychiatry. https://doi.org/10.1001/jamapsychiatry.2023.2282