# 2020 Decennial Census Effect on the National Survey on Drug Use and Health (NSDUH) Substance Use and Mental Health Estimates

Kathy Spagnola[1], Neeraja Sathe[1], Patrick Chen[1],
Devon Cribb[1], Rong Cai[2], Jingsheng Yan[2], Jennifer M. Hoenig[2]
[1]RTI International, 3040 Cornwallis Road, RTP, NC 27709
[2]SAMHSA, Center for Behavioral Health Statistics and Quality, Rockville, MD 20857

**Abstract**
The National Survey on Drug Use and Health (NSDUH) provides national estimates of substance use and mental health among the civilian, noninstitutionalized population aged 12 or older in the United States. NSDUH person-level analysis weights are calibrated to estimated totals of the target population to reduce coverage bias and variance of survey estimates. The U.S. Census Bureau produces population estimates annually, using current data on births, deaths, and migration to adjust the most recent decennial census data. These postcensal population estimates become less accurate with each passing year. The person-level analysis weights used in the 2020 NSDUH estimates were calibrated to the 2020 postcensal population estimates based on the 2010 decennial census. In the 2021 NSDUH, population estimates began using the 2020 decennial census in the person-level analysis weights. The 2020 decennial census effect study examined whether, and to what extent, 2020 NSDUH estimates would have differed if person-level weights were calibrated to the 2020 decennial census data instead of the postcensal population estimates anchored to the 2010 decennial census.

**Key Words:** 2020 population, census effect, NSDUH, weighting, substance use, mental health

## 1. Overview/Study Background

National Survey on Drug and Health (NSDUH) is an annual national survey that provides national representative data on the use of tobacco, alcohol, and illicit drugs; substance use disorders; mental health issues; and the use of mental health services among the civilian, noninstitutionalized population aged 12 or older in the United States. NSDUH has a stratified five-stage cluster design (Center for Behavioral Health Statistics and Quality, 2021a) with a target of 67,500 respondents. In a typical year, 220,000 households are screened and 67,500 interviews are completed. However, because of coronavirus disease 2019 (COVID-19), NSDUH in-person data collection was suspended on March 16, 2020. Ongoing COVID-19 infection rates in the United States made it nearly impossible to perform conventional in-person data collection, which could reduce the respondent sample size to an unacceptable level. Therefore, SAMHSA approved multimode data collection (in person and via the web) for the 2020 NSDUH beginning in Quarter 4 (October to December) of 2020. In-person data collection resumed on October 1, 2020 (in locations where COVID-19 infection metrics were sufficiently low), and web-based data collection began on October 30, 2020. The multimode data collection continued to be used in the 2021, 2022, and 2023 NSDUHs and will be implemented in future NSDUHs.

The analysis weights for NSDUH are developed to obtain unbiased estimates and are benchmarked to the national target population counts for various demographic (age, gender, race/ethnicity) and geographic domains (state) in the last poststratification adjustment to reduce coverage bias and variance of survey estimates. In the 2020 and 2021 NSDUHs, the analysis weights were based on population estimates from different decennial censuses. Specifically, the analysis weights used in producing the published 2020 NSDUH estimates were poststratified to the population estimates for 2020 based on the 2010 decennial census. Starting with the 2021 NSDUH, population estimates based on the 2020 decennial census were used in developing the person-level analysis weights. This paper discusses the 2020 census effect study (CES), which examined the effect of the differences between the 2020 population estimates from the 2020 census and the 2020 population estimates from the 2010 census on the 2020 NSDUH estimates. The purpose of the CES was to determine whether a decennial census effect existed and, if so, the magnitude of the effect. Because the use of multimode data collection (in person and via the web) throughout 2021 affected the comparability of estimates between 2020 and 2021, the 2020 census effect investigation focused only on the effects of the population estimates from the 2010 and 2020 censuses on substance use and mental health estimates for the 2020 NSDUH. The 2020 census effect investigation did not evaluate decennial census effects on trends (or changes) in estimates between 2020 and 2021.

## 1.1 Comparison of 2010 and 2020 Census-Based Target Population Estimates for 2020

The 2020 population estimates for the civilian, noninstitutionalized population of the United States aged 12 or older residing within the 50 states and the District of Columbia based on the 2010 decennial census and the 2020 decennial census were similar with a difference of 0.69% (i.e., the population estimate based on the 2020 census was 0.69% higher than the population estimate based on the 2010 census), as shown in Table 1. The 2010 decennial census-based population estimates underestimated population counts in all demographic domains listed in Table 1, except for people aged 26 to 34 and individuals aged 65 or older. For those two age groups, the population estimate based on the 2020 census was 0.62% and 2.07% lower than the population estimate based on the 2010 census, respectively. Among all age groups, people aged 12 to 17 had the largest underestimate (4.1%); males were underestimated (1.39%) more than females; and among the race groups, those identifying as Native Hawaiian or Other Pacific Islander were underestimated the most (2.16%). Hispanic or Latino people were more underestimated (1.03%) than those identifying as non-Hispanic or Latino.

**Table 1:** Differences between the 2020 Civilian Noninstitutionalized Population Counts for People Aged 12 Years or Older Based on the 2010 and 2020 Decennial Censuses, by Age, Gender, Hispanic Origin, and Race

| Domain | 2020 Population Based on 2010 Census | 2020 Population Based on 2020 Census | Percentage Difference Relative to 2020 Population Based on 2010 Census[1] |
|---|---|---|---|
| *Total* | 276,889,195 | 278,792,982 | 0.69% |
| 12 to 17 | 24,994,109 | 26,019,594 | 4.10% |
| 18 to 25 | 33,497,220 | 33,518,069 | 0.06% |
| 26 to 34 | 40,537,211 | 40,287,600 | -0.62% |
| 35 to 49 | 61,014,613 | 62,170,356 | 1.89% |
| 50 to 64 | 62,350,083 | 63,431,418 | 1.73% |
| 65 or Older | 54,495,960 | 53,365,945 | -2.07% |
| Male | 134,379,963 | 136,252,696 | 1.39% |
| Female | 142,509,232 | 142,540,287 | 0.02% |
| Hispanic or Latino | 48,381,323 | 48,881,321 | 1.03% |
| Not Hispanic or Latino | 228,507,872 | 229,911,661 | 0.61% |
| White[2] | 213,372,733 | 214,712,160 | 0.63% |
| Black or African American[2] | 35,861,973 | 36,185,815 | 0.90% |
| AIAN[2] | 3,395,628 | 3,426,682 | 0.91% |
| Asian[2] | 17,105,448 | 17,238,450 | 0.78% |
| NHOPI[2] | 659,267 | 673,525 | 2.16% |
| Two or More Races[2] | 6,494,147 | 6,556,349 | 0.96% |

AIAN = American Indian or Alaska Native; NHOPI = Native Hawaiian or Other Pacific Islander.
Note: Population counts are annualized estimates (averaged over four quarters) of the 2020 population and reflect the population of the entire year.
[1] Based on the formula {[(2020 Population Based on 2020 Census) − (2020 Population Based on 2010 Census)] ÷ (2020 Population Based on 2010 Census)} × 100.
[2] Race domains in this table include Hispanic or Latino in addition to people who are not Hispanic or Latino.
Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, 2020.

## 1.2 Comparison of Two Sets of Analysis Weights for 2020 NSDUH

The analysis weights for the 2020 NSDUH were the product of 16 weight components for the selection of sampling units at different stages, nonresponse, poststratification, and extreme weight adjustments at the dwelling unit or person level. The person-level poststratification adjustment calibrated nonresponse-adjusted weights to the target population estimates. The target population estimates in the person-level poststratification adjustment for the analysis weights used for published 2020 estimates were based on 2020 population estimates from the 2010 decennial census. Because of the impact of COVID-19, data collection was interrupted in Quarter 2 and Quarter 3 of 2020. Data collection in Quarter 1 was in person, and multimode data collection (in person and via the web) was introduced in Quarter 4. Thus, quarterly analysis weights for Quarter 1 and Quarter 4 were developed, and the annual analysis weights were calculated by averaging the Quarter 1 and Quarter 4 analysis weights (Center for Behavioral Health Statistics and Quality, 2022a).

For the CES, a new set of analysis weights for the 2020 NSDUH was developed using the population estimates based on the 2020 decennial census as control totals in the final poststratification adjustment. The same procedures for developing the original analysis weights were followed. The poststratification adjustment was conducted separately for Quarter 1 and Quarter 4, and the annual analysis weights were the average of the analysis weights for Quarter 1 and Quarter 4 (Center for Behavioral Health Statistics and Quality, 2022a). This new set of analysis weights for the 2020 NSDUH (new weights) was created and compared with the 2020 original analysis weights (old weights).

At the national level, the distributions and unequal weighting effect (UWE) of old weights and new weights were very similar (Table 2). For demographic domains, no large changes in the distributions and UWE were observed.

**Table 2:** Distribution Comparison between the New Weights and Old Weights for 2020

| Weights | n | Mean | Min | p25 | Median | p75 | Max | UWE |
|---------|-----|------|-----|------|--------|------|---------|------|
| Old | 36,284 | 7,632 | 3 | 1,258 | 3,636 | 9,167 | 216,420 | 3.31 |
| New | 36,284 | 7,685 | 3 | 1,283 | 3,671 | 9,240 | 234,101 | 3.31 |

Max = maximum; min = minimum; p25 = 25th percentile; p75 = 75th percentile; UWE = unequal weighting effect. UWE refers to the contribution in the design effect due to unequal selection probability and is defined as $1 + \left[(n-1)/n\right] * CV^2$, where $CV$ = (coefficient of variation of weights)/100, and $n$ is the sample size.

Note: Old weights are the original analysis weights calibrated to the control totals based on the 2010 decennial census. New weights are the new analysis weights calibrated to the control totals based on the 2020 decennial census.

## 1.3 Methodology of the CES

Two sets of analysis weights are available for the 2020 NSDUH, the original analysis weights (referred to as "old weights") and the newly developed analysis weights (referred to as "new weights") for the CES. Throughout this paper, estimates (either estimated percentages or totals) from the 2020 NSDUH published tables based on the original analysis weights are referred to as "2020(Old)." Estimates generated using the new weights are referred to as "2020(New)."

The census effect term is defined as the difference between 2020(New) and 2020(Old) estimates. There are two ways to view the census effect:

1. **Comparing 2020(New) and 2020(Old) Estimates.** This comparison includes performing statistical tests for differences in estimated percentages and totals. The null hypothesis is testing the difference [2020(New) − 2020(Old) = 0], which is equivalent to testing the ratio [2020(New) / 2020(Old) = 1].[1] Ratios close to 1 would indicate a minimal census effect, and ratios further away from 1 would indicate a larger census effect. Statistical testing results from the estimated differences can be used to determine whether the estimated ratios

---

[1] Technically, when testing estimated differences and testing estimated ratios, it is assumed that the ratios and differences of estimates are asymptotically normal, a common and reasonable assumption when dealing with data from a complex survey sample, such as NSDUH.

(for the percentages and numbers in thousands) are significantly different from 1. Tests of differences are discussed in this paper.

2. **Comparing Changes in Statistical Significance (*p* values).** This comparison includes conducting significance testing for 2021 versus 2020(Old) and 2021 versus 2020(New). Comparing the results from each can determine whether the significance between 2021 and 2020 estimates was dependent on the weight used for 2020 (Old or New). Many changes in the significance results would indicate the existence of a census effect.

Because of COVID-19, multimode data collection was introduced in the 2020 NSDUH in October 2020 and was used throughout the 2021 NSDUH. For multimode data collection, respondents completed the survey via the web or in person in eligible locations. Additionally, the 2020 NSDUH estimates were based on only two quarters of data (because data were not collected in Quarters 2 and 3). Methodological investigations led to the conclusion that estimates based on multimode data collection in 2021 were not comparable with estimates from 2020 because of the effects of missing some quarters and because of different data collection modes (both years were multimode but 2020 quarter 1 only had in person data collection). Therefore, this CES focused only on direct comparisons between 2020(New) and 2020(Old) estimates (number 1 in the list above) to assess the census effect. Thus, estimates between 2020 and 2021 were not compared as part of this CES.

## 2. Results and Findings

Sixty-six tables of substance use and mental health estimates from the *Results from the 2020 National Survey on Drug Use and Health: Detailed Tables* (Center for Behavioral Health Statistics and Quality, 2021b) were selected to replicate using the new analysis weights based on the 2020 census control totals. These tables were used to determine the impact of using the 2020 census-based control totals. The substance use tables were chosen to represent a diverse group of substance use categories, such as illicit drug use, alcohol use, cigarette use, initiation of substances, mean age at first use of substances, perceived risk of substance use, substance use disorder, and treatment. The mental health tables cover a range of topics, such as mental illness, major depressive episode, and mental health treatment. Both sets of tables show these measures by various demographic, socioeconomic, and geographic subgroups, such as age, race/ethnicity, gender, employment, region, county type, poverty level, education, and health insurance coverage.

In this paper, summary results from the 66 tables described above are discussed. The substance use tables and mental health tables were summarized separately. These summary results include looking at the ratios of the new estimate divided by the old estimate. Significance testing between the 2020(New) and 2020(Old) estimates was conducted using a *t*-test (with the appropriate degrees of freedom). Lastly, for estimates that were significantly different, absolute relative differences (100 times the absolute value of the old estimate minus the new estimate divided by the old estimate) were calculated. The next three sections discuss these summary results.

### 2.1 Ratio of Estimates

All ratios of estimates and standard errors in the selected tables for substance use measures and mental health measures were calculated using these equations:

Ratio of Estimate = 2020(New) Estimate / 2020(Old) Estimate

Ratio of Standard Error (SE) = 2020(New) SE / 2020(Old) SE

Unrounded estimates (for percentages and totals) were used in the calculation of these ratios, then the ratios were rounded to two decimal places in the tables. If the ratios of the estimates and the standard errors were close to 1, then the two estimates were considered to be similar. For example, the 2020(Old) and 2020(New) estimated percentages of past month ecstasy use among people aged 12 or older were 0.2% and 0.3%, respectively, which yielded a ratio of the estimates that was 1.02 (the ratio was calculated using unrounded percentages) (Table 4). Testing the null hypothesis that the ratios of estimates are equal to 1 is equivalent to testing the null hypothesis that the differences of estimates are equal to 0.

Distributions of the ratios of the estimates for substance use and mental health measures are summarized in Table 5. Additionally, *t*-tests were performed to determine whether the mean of the ratios (for all estimate types) were significantly different from 1. Among substance use measures, the ratios of the estimated percentages ranged from 0.75 to 1.43, with a mean and median of 1.00; the mean of the ratios of the estimated percentages was not significantly different from 1. The ratios of the estimated totals were similar, ranging from 0.76 to 1.44, with a mean and median of 1.01; however, the mean was significantly different from 1 at the 0.05 significance level. The maximum ratio was observed for past month PCP use among people aged 12 or older for both percentages and totals. Past month PCP use is a very low prevalence outcome. For past month PCP use among peopled aged 12 or older, the estimated percentages for both 2020(Old) and 2020(New) were 0.0% for people aged 12 or older (the totals were 24,000 and 34,000, respectively) (Tables 3 and 4).

The mean and median ratios for percentages and totals for the mental health outcomes (Table 5) were close to 1, but the maximum ratios (1.78 for the percentages and 1.91 for the totals) were higher than those for the substance use measures. These mean ratios for the estimated mental health percentages and totals were significantly different from 1 at the 0.05 significance level.

For both the substance use and mental health outcomes, the ratio of the standard errors of estimated percentages and totals followed a similar pattern, but the range was slightly wider than the range of the ratio of the estimates.

**Table 3:** Types of Illicit Drug Use in Lifetime and Past Month: Among People Aged 12 or Older; Numbers in Thousands, 2020

| Time Period/Drug | 2020(Old) | 2020(New) | Ratio of Estimates |
|---|---|---|---|
| *Lifetime* | | | |
| Illicit Drugs[1] | 138,543[b] | 139,393 | 1.01 |
|   Marijuana | 126,504[b] | 127,390 | 1.01 |
|   Cocaine | 39,261[b] | 39,797 | 1.01 |
|   Heroin | 6,252 | 6,323 | 1.01 |
|   Hallucinogens | 43,949[b] | 44,307 | 1.01 |
|     LSD | 28,123[b] | 28,402 | 1.01 |
|     PCP | 6,141 | 6,224 | 1.01 |
|     Ecstasy | 20,478[b] | 20,703 | 1.01 |
| *Past Month* | | | |
| Illicit Drugs[1] | 37,309[a] | 37,641 | 1.01 |
|   Marijuana | 32,784[a] | 33,026 | 1.01 |
|   Cocaine | 1,831 | 1,948 | 1.06 |
|   Heroin | 513 | 518 | 1.01 |
|   Hallucinogens | 1,761 | 1,794 | 1.02 |
|     LSD | 649[a] | 686 | 1.06 |
|     PCP | 24 | 34 | 1.44 |
|     Ecstasy | 681 | 700 | 1.03 |

Note: The 2020(Old) and 2020(New) estimates were generated using different analysis weights. The weights were poststratified using 2020 population estimates based on the 2010 census for 2020(Old) and on the 2020 census for 2020(New).

Note: Ratio of Estimate = 2020(New) Estimate / 2020(Old) Estimate.

[a,b] Difference between 2020(Old) estimate and 2020(New) estimate is statistically significant at the 0.05 level (a) or 0.01 level (b).

[1] Illicit Drug Use in Lifetime includes the misuse of prescription psychotherapeutics, even though those estimates are not reported due to potential underreporting in lifetime prescription psychotherapeutics estimates.

Definitions: Measures and terms are defined in Appendix A of the *Results from the 2020 National Survey on Drug Use and Health: Detailed Tables* at https://www.samhsa.gov/data/report/2020-nsduh-detailed-tables.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, Quarters 1 and 4, 2020.

**Table 4:** Types of Illicit Drug Use in Lifetime and Past Month:
Among People Aged 12 or Older; Percentages, 2020

| Time Period/Drug | 2020(Old) | 2020(New) | Ratio of Estimates |
|---|---|---|---|
| *Lifetime* | | | |
| Illicit Drugs[1] | 50.0 | 50.0 | 1.00 |
|   Marijuana | 45.7 | 45.7 | 1.00 |
|   Cocaine | 14.2[a] | 14.3 | 1.01 |
|   Heroin | 2.3 | 2.3 | 1.00 |
|   Hallucinogens | 15.9 | 15.9 | 1.00 |
|     LSD | 10.2 | 10.2 | 1.00 |
|     PCP | 2.2 | 2.2 | 1.01 |
|     Ecstasy | 7.4 | 7.4 | 1.00 |
| *Past Month* | | | |
| Illicit Drugs[1] | 13.5 | 13.5 | 1.00 |
|   Marijuana | 11.8 | 11.8 | 1.00 |
|   Cocaine | 0.7 | 0.7 | 1.06 |
|   Heroin | 0.2 | 0.2 | 1.00 |
|   Hallucinogens | 0.6 | 0.6 | 1.01 |
|     LSD | 0.2 | 0.2 | 1.05 |
|     PCP | 0.0 | 0.0 | 1.43 |
|     Ecstasy | 0.2 | 0.3 | 1.02 |

Note: The 2020(Old) and 2020(New) estimates were generated using different analysis weights. The weights were poststratified using 2020 population estimates based on the 2010 census for 2020(Old) and on the 2020 census for 2020(New).

Note: Ratio of Estimate = 2020(New) Estimate / 2020(Old) Estimate.

[a,b] Difference between 2020(Old) estimate and 2020(New) estimate is statistically significant at the 0.05 level (a) or 0.01 level (b).

[1] Illicit Drug Use in Lifetime includes the misuse of prescription psychotherapeutics, even though those estimates are not reported due to potential underreporting in lifetime prescription psychotherapeutics estimates.

Definitions: Measures and terms are defined in Appendix A of the *Results from the 2020 National Survey on Drug Use and Health: Detailed Tables* at https://www.samhsa.gov/data/report/2020-nsduh-detailed-tables.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, Quarters 1 and 4, 2020.

**Table 5:** Distribution of the Ratios of Estimated Percentages and Estimated Totals for Substance Use and Mental Health Outcomes

| Summary Statistics | Substance Use | | Mental Health | |
| | Ratios of Estimated Percentages ($n = 715$) | Ratios of Estimated Totals ($n = 685$) | Ratios of Estimated Percentages ($n = 1,086$) | Ratios of Estimated Totals ($n = 1,086$) |
| --- | --- | --- | --- | --- |
| Min. | 0.75 | 0.76 | 0.80 | 0.80 |
| 1% | 0.92 | 0.92 | 0.89 | 0.93 |
| 5% | 0.97 | 0.98 | 0.96 | 0.97 |
| 10% | 0.98 | 0.99 | 0.97 | 0.98 |
| 25% | 0.99 | 1.00 | 0.99 | 1.00 |
| Median | 1.00 | 1.01 | 1.00 | 1.01 |
| 75% | 1.01 | 1.02 | 1.01 | 1.03 |
| 90% | 1.02 | 1.04 | 1.03 | 1.06 |
| 95% | 1.04 | 1.06 | 1.06 | 1.07 |
| 99% | 1.09 | 1.12 | 1.18 | 1.20 |
| Max | 1.43 | 1.44 | 1.78 | 1.91 |
| Mean | 1.00 | 1.01 | 1.00 | 1.02 |
| Max./Mean | 1.43 | 1.42 | 1.77 | 1.87 |
| Mean/Min. | 1.33 | 1.33 | 1.25 | 1.27 |
| CV | 3.62 | 3.98 | 5.02 | 5.33 |

CV = coefficient of variation; max. = maximum; min. = minimum.
Note: Ratio of Estimate = 2020(New) Estimate / 2020(Old) Estimate.
Note: The 2020(Old) and 2020(New) estimates were generated using different analysis weights. The weights were poststratified using 2020 population estimates based on the 2010 census for 2020(Old) and on the 2020 census for 2020(New).
Note: Estimated totals are numbers in thousands.
Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, Quarters 1 and 4, 2020.

## 2.2 Testing of 2020(Old) versus 2020(New) Estimates

Significance testing between the 2020(New) and 2020(Old) estimates were conducted. When comparing prevalence estimates, the null hypothesis (no difference between prevalence estimates) was tested against the alternative hypothesis (there was a difference in prevalence estimates) using a *t*-test (with the appropriate degrees of freedom).[2] For more details about these tests, see Section 3.2.3 of the *2021 National Survey on Drug Use and Health (NSDUH): Methodological Summary and Definitions* (Center for Behavioral Health Statistics and Quality, 2022b). Throughout the rest of this section, a significant increase refers to cases where the 2020(New) estimate was significantly higher than the 2020(Old) estimate at the 5% level of significance, and vice versa for significant decreases.

Table 6 shows the numbers (and percentages) of significant differences (increases and decreases) between the 2020(Old) and 2020(New) estimates among substance use

---

[2] Testing whether estimated ratios of the percentages and numbers in thousands are significantly different from 1 is equivalent to testing whether estimated differences in the percentages and numbers in thousands are significantly different from 0. Results from the *t* tests of differences in the 2020(New) and 2020(Old) estimated percentages and numbers in thousands are discussed in this section.

outcomes by each outcome type. For estimated percentages for all substance use outcomes, among 715 pairs of estimated percentages that were tested, there were 23 cases (3.2% of total) with a significant increase and 46 cases (6.4%) with a significant decrease. For totals, of the 685 sets of estimates that were tested, the new estimated total was significantly higher than the old total 218 times (31.8% of all pairs), and 2020(New) was significantly lower than 2020(Old) in only 12 cases (1.8%).

These findings indicate that the estimated totals were more likely to be impacted by using weights based on 2020 census population estimates from the 2020 census instead of the 2010 census. This was likely because changes in weights affected the numerator and denominator somewhat proportionally while estimating the percentages.

For lifetime cocaine use among people aged 12 or older, the estimated percentage and total increased from 2020(Old) to 2020(New). Specifically, the 2020(Old) estimate was 39.3 million people (14.2%), and the 2020(New) estimate was 39.8 million (14.3%) (Tables 3 and 4). For both the totals and percentages, the 2020(New) and 2020(Old) estimates were significantly different. This example shows that the estimates (both percentages and totals) can be significantly different, even if the actual numeric difference does not seem that large. Because the sample sizes of the two domains being tested can be large at times, and because the underlying data were the same, minor differences can be flagged as being significant.

**Table 6:** Counts of the Significant Differences (at the 5 Percent Level of Significance) and Absolute Relative Differences in Estimated Percentages and Estimated Totals, by Substance Use Outcome Type

| Estimate Type/ Outcome Type | Number of Estimates | Sig. Increases (%)[1,2] | Median ARD[2,3] among Increases | Max. ARD[2,3] among Increases | Sig. Decreases (%)[1,2] | Median ARD[2,3] among Decreases | Max. ARD[2,3] among Decreases |
|---|---|---|---|---|---|---|---|
| Estimated Percentages | | | | | | | |
| Overall | 715 | 23 (3.2) | 0.76 | 6.07 | 46 (6.4) | 1.55 | 7.34 |
| Illicit Drugs | 193 | 4 (2.1) | 0.78 | 0.99 | 3 (1.6) | 2.95 | 4.00 |
| Tobacco | 137 | 9 (6.6) | 0.68 | 6.07 | 11 (8.0) | 0.57 | 7.34 |
| Alcohol | 135 | 6 (4.4) | 0.91 | 3.38 | 24 (17.8) | 1.60 | 3.53 |
| SUD/Treatment | 123 | 0 (0.0) | n/a | n/a | 0 (0.0) | n/a | n/a |
| Mean Age | 49 | 0 (0.0) | n/a | n/a | 7 (14.3) | 0.54 | 0.91 |
| Other[4] | 78 | 4 (5.1) | 0.79 | 3.60 | 1 (1.3) | 1.79 | 1.79 |
| Estimated Totals | | | | | | | |
| Overall | 685 | 218 (31.8) | 1.30 | 8.97 | 12 (1.8) | 0.58 | 6.76 |
| Illicit Drugs | 193 | 48 (24.9) | 1.11 | 7.96 | 1 (0.5) | 3.94 | 3.94 |
| Tobacco | 137 | 48 (35.0) | 1.01 | 7.25 | 4 (2.9) | 1.61 | 6.76 |
| Alcohol | 135 | 42 (31.1) | 1.04 | 4.72 | 7 (5.2) | 0.53 | 2.41 |
| SUD/Treatment | 123 | 36 (29.3) | 2.72 | 6.12 | 0 (0.0) | n/a | n/a |
| Initiation | 19 | 2 (10.5) | 2.11 | 2.40 | 0 (0.0) | n/a | n/a |
| Other[4] | 78 | 42 (53.8) | 4.06 | 8.97 | 0 (0.0) | n/a | n/a |

ARD = absolute relative difference; max. = maximum; n/a = not applicable; sig. = significant; SUD = substance use disorder.

Note: Unrounded estimates were used to determine increases and decreases.

Note: Estimated totals are numbers in thousands.

[1] Number of significant differences (increases or decreases) at 0.05 level. When the 2020(New) estimate is significantly greater than the 2020(Old) estimate, it indicates an increase, and when the 2020(New) estimate is less than the 2020(Old) estimate, it indicates a decrease.

[2] The 2020(Old) and 2020(New) estimates were generated using different analysis weights. The weights were poststratified using 2020 population estimates based on the 2010 census for 2020(Old) and on the 2020 census for 2020(New).

[3] ARD = 100*abs[2020(Old) - 2020(New)] / 2020(Old). ARDs were calculated only for estimates that were significantly different.

[4] Other outcome group includes driving under the influence and perceived great risk of using a substance.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, Quarters 1 and 4, 2020.

Table 7 shows the comparisons between 2020(New) and 2020(Old) for estimated percentages and estimated totals by age group for mental health outcomes. For example, the "18 or older" row includes all estimates among that age group including people aged 18 or older crossed with other domains. The "18 or older" row does not include estimates among age subgroups (like people aged 18 to 25).

A large number of significant increases for estimated totals (54.1% for the mental health outcomes) were seen for youths aged 12 to 17 (Table 7). These results seem to make sense given that those aged 12 to 17 had the largest difference between the 2020 population based on the 2010 and 2020 census (specifically, the relative difference was 4.10%, nearly twice the difference of the next largest difference among age groups, gender, Hispanic origin, and race) (Table 1).

**Table 7:** Counts of the Significant Differences (at the 5 Percent Level of Significance) and Absolute Relative Differences in Mental Health Estimated Percentages and Estimated Totals, by Age Group

| Estimate Type/Age Group | Number of Estimates | Sig. Increases (%)[1,2] | Median ARD[2,3] among Increases | Max. ARD[2,3] among Increases | Sig. Decreases (%)[1,2] | Median ARD[2,3] among Decreases | Max. ARD[2,3] among Decreases |
|---|---|---|---|---|---|---|---|
| **Estimated Percentages** | | | | | | | |
| Overall MH | 1086 | 17 (1.6) | 2.23 | 7.97 | 46 (4.2) | 1.93 | 5.68 |
| 18 or Older | 521 | 13 (2.5) | 1.50 | 7.97 | 21 (4.0) | 1.74 | 5.68 |
| 18 to 25 | 80 | 0 (0.0) | n/a | n/a | 5 (6.3) | 2.24 | 3.81 |
| 26 or Older | 10 | 0 (0.0) | n/a | n/a | 0 (0.0) | n/a | n/a |
| 26 to 49 | 79 | 1 (1.3) | 3.99 | 3.99 | 3 (3.8) | 1.56 | 2.04 |
| 50 or Older | 78 | 1 (1.3) | 2.73 | 2.73 | 1 (1.3) | 2.03 | 2.03 |
| 12 to 17 | 318 | 2 (0.6) | 3.13 | 3.37 | 16 (5.0) | 2.88 | 5.23 |
| **Estimated Totals** | | | | | | | |
| Overall MH | 1086 | 216 (19.9) | 3.89 | 9.29 | 17 (1.6) | 1.80 | 4.25 |
| 18 or Older | 521 | 24 (4.6) | 1.87 | 8.45 | 9 (1.7) | 2.06 | 3.64 |
| 18 to 25 | 80 | 1 (1.3) | 4.16 | 4.16 | 5 (6.3) | 2.04 | 4.25 |
| 26 or Older | 10 | 0 (0.0) | n/a | n/a | 0 (0.0) | n/a | n/a |
| 26 to 49 | 79 | 19 (24.1) | 1.10 | 5.01 | 0 (0.0) | n/a | n/a |
| 50 or Older | 78 | 0 (0.0) | n/a | n/a | 3 (3.8) | 1.75 | 1.80 |
| 12 to 17 | 318 | 172 (54.1) | 4.13 | 9.29 | 0 (0.0) | n/a | n/a |

ARD = absolute relative difference; max. = maximum; MH = mental health; n/a = not applicable; sig. = significant.

Note: Unrounded estimates were used to determine increases and decreases.

Note: Estimated totals are numbers in thousands.

Note: There were several age groups that had 10 or fewer estimates that were lumped in with other age groups. Specifically, the 12 to 13, 14 to 15, and 16 to17 age groups were included in the 12 to 17 row.

Note: Each row is mutually exclusive, and the total estimate column sums to the total number of mental health estimate pairs.

[1] Number of significant differences (increases or decreases) at 0.05 level. When the 2020(New) estimate is significantly greater than the 2020(Old) estimate, it indicates an increase, and when the 2020(New) estimate is less than the 2020(Old) estimate, it indicates a decrease.

[2] The 2020(Old) and 2020(New) estimates were generated using different analysis weights. The weights were poststratified using 2020 population estimates based on the 2010 census for 2020(Old) and on the 2020 census for 2020(New).

[3] ARD = 100*abs[2020(Old) - 2020(New)] / 2020(Old). ARDs were calculated only for estimates that were significantly different.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, Quarters 1 and 4, 2020.

## 2.3 Absolute Relative Differences

In addition to showing the number of significant differences (discussed in Section 2.2), Tables 6 and 7 also show the distribution of the absolute relative differences (ARDs) among the estimates that were significantly different. The formula for calculating the ARD is as follows:

$$ARD = 100*abs(t1 - t2)/t1, \text{ where}$$

t1 is the 2020 NSDUH published estimate [2020(Old)], and

t2 is the 2020(New) estimate.

For the substance use estimates, the maximum ARD was 7.34 for the estimated percentages and 8.97 for the estimated totals. The maximum ARD for estimated percentages was among the tobacco outcome group. The maximum ARD for the estimated totals was among the "other" outcome group, which includes perceived risk of substance use and driving under the influence outcomes. For percentages that increased, the overall median ARD was 0.76; among decreases, the median ARD was 1.55. For totals, the median ARD was 1.30 and 0.58, respectively, for increases and decreases (Table 6).

Among the increases for the mental health estimates, the maximum ARD was 7.97 for the estimated percentages and 9.29 for the estimated totals. The maximum ARD for estimated percentages was among people aged 18 or older. The maximum ARD for the estimated totals was among the youth (those aged 12 to 17). For percentages that increased, the overall median ARD was 2.23; among decreases, the median ARD was 1.93. For totals, the median ARD was 3.89 and 1.80, respectively, for increases and decreases. (Table 7).

## 3. Conclusion

For substance use, 9.6% of estimated percentages were statistically different between 2020(New) and 2020(Old) at the 0.05 significance level, whereas 33.6% of estimated totals were statistically different (Table 6). For mental health estimates, 5.8% of estimated percentages and 21.5% of estimated totals were statistically different between 2020(New) and 2020(Old) (Table7). This indicates that the census effect exists in 2020 estimates; however, mental health estimates were less influenced by the census effect than substance use estimates. The census effect had less impact on estimated percentages than estimated totals for both substance use and mental health. Although the proportions of significant results were high, particularly for the estimated totals, the magnitude of the census effect should be relatively low given the following:

- Trivial differences can be flagged as statistically significant in direct comparisons between 2020(New) and 2020(Old) because the underlying data were the same and sample sizes were large in most comparison groups.
- The mean and median of ratios [2020(New) / 2020(Old)] for estimated percentages and total numbers were close to 1, suggesting that 2020(New) estimates were similar to 2020(Old) estimates (Table 5).
- Among estimates that were statistically different between 2020(New) and 2020(Old), the relative differences tended to be small. The median absolute relative difference was less than 2% for overall substance use and less than 4% for overall mental health. The maximum ARD was less than 10% (Tables 6 and 7).

As mentioned earlier in this paper, this study also examined the potential for census effects on age group. For estimated totals of substance use, significant differences were more concentrated in people aged 12 to 17. For mental health estimates, a higher proportion of significant results was observed for estimated totals in people aged 12 to 17. This suggests that the census effect was more of a concern in the 12 to 17 age group than in other age groups.

## Acknowledgment

## Disclaimer

The views expressed in this document do not necessarily reflect the official policies of the U.S. Department of Health and Human Services, nor does mention of trade names, commercial practices, or organizations imply endorsement by the U.S. Government.

## References

Center for Behavioral Health Statistics and Quality. (2021a). *2020 National Survey on Drug Use and Health (NSDUH) methodological resource book, Section 2: Sample design report*. https://www.samhsa.gov/data/report/nsduh-2020-sample-design-report

Center for Behavioral Health Statistics and Quality. (2021b). *Results from the 2020 National Survey on Drug Use and Health: Detailed tables*. https://www.samhsa.gov/data/report/2020-nsduh-detailed-tables

Center for Behavioral Health Statistics and Quality. (2022a). 2020 National Survey on Drug Use and Health (NSDUH): Methodological resource book, Section 11: Person-level sampling weight calibration. https://www.samhsa.gov/data/report/nsduh-2020-person-level-sampling-weight-calibration-report

Center for Behavioral Health Statistics and Quality. (2022b). *2021 National Survey on Drug Use and Health (NSDUH) methodological summary and definitions*. https://www.samhsa.gov/data/report/2021-methodological-summary-and-definitions