Tobacco Use Supplement
Current Population Survey (TUS-CPS)

2020 Informational Session Questions and Responses

The National Cancer Institute (NCI) TUS-CPS team has compiled responses to questions submitted by registrants for the 2020 Society for Research on Nicotine and Tobacco (SRNT) TUS-CPS Informational Session, though the session was cancelled. Please find all questions and responses within this document. Use the Table of Contents below to access the various topic areas.

For additional information about TUS-CPS, please visit cancercontrol.cancer.gov/tus-cps.

For specific questions, please contact the NCI TUS-CPS team at ncidccpsbrpadvances@mail.nih.gov.

Table of Contents

TUS Survey Design and Data

1. **Has the survey design changed over the years?**
   While the basic approach to collecting data has remained intact, improvements to CPS survey design and methodology have been made over time to ensure the use of best practices and collection of high-quality data. Some recent sampling changes include moving from once-per-decade to annual sampling and the adoption of the Master Address File (MAF), featuring the most updated and extensive address information from the United States Postal Service.

   Historically, core TUS-CPS survey questions have remained the same, but some recent versions have featured new questions. For example, some versions have had a special focus on tobacco cessation and/or menthol cigarette use. The 2018-2019 TUS-CPS included new questions categorizing e-cigarette device type; policies about e-cigarette use in the workplace and rules inside the home; and exposure to federal court-mandated tobacco company statements about the dangers of smoking cigarettes. The new topics added to recent survey waves are highlighted here: https://cancercontrol.cancer.gov/brp/tcrb/tus-cps/surveytopics.html. In addition to changes in questions, there have also been modifications to respondent selection criteria.

   For more details about survey design and methodology, please refer to the CPS *Technical Paper 77, Design and Methodology*, available at https://www.census.gov/programs-surveys/cps/technical-documentation/complete.html. For additional questions, please contact Tim Marshall of the Census Bureau's Demographic Surveys Division at (301) 763-3806.

2. **Is it possible to account for the complex survey design when analyzing TUS data? My understanding is that TUS does not provide the necessary cluster/stratum variable or replicate weights.**
   CPS and other supplement files, including replicate weight files (2010-2019), are available on the CPS FTP site. TUS-CPS does not contain cluster/stratum variables and instead uses replicate weights for a more robust variance estimation to calculate confidence intervals. For additional details on specifying the replicate weights using the Fay method, please refer to the TUS-CPS Technical Documentation for each survey wave (Appendix 16–Source and Accuracy Statement). The replicate weights are available on the Census FTP site. The webinar linked here should also be helpful for describing how the replicate weights were derived (http://www.cancermeetings.org/TUSCPSWebinar/documents/tuscps_Webinar_Liu.pdf).

3. **Why are some individuals sampled twice in different months of each wave of the TUS survey? How should we approach obtaining prevalence estimates accounting for the fact that we are double-counting some individuals?**
   Respondents are typically included in the 2018-19 TUS-CPS sample only once. The rotation of households in and out of sample for the TUS-CPS was generally sufficiently spaced with months between interviewing, so the sample did not contain many overlapping panels. This eliminated most individuals being in the sample twice within the 2018-19 main data files. The overlap is even smaller if you are working with self-responses only. Some people on overlapping panels took TUS by proxy the first time and self-responded the second time or vice versa; some took TUS by proxy both times. Thus, the small error introduced by the overlapping respondents can be ignored for most purposes. Additional information about sampling can be found in the technical documentation for each TUS survey wave, available at https://cancercontrol.cancer.gov/brp/tcrb/tus-cps/questionnaires.html.

4. **Is there any oversampling issue in data collection? If so, how was this issue handled?**
The TUS-CPS is a periodic supplement of the CPS conducted by Census Bureau staff. The TUS follows the CPS's state-based sampling design, with the sample in each state being independent of the others. As a result of the sample being selected from multiple frames using multiple stages of selection, the TUS-CPS does not oversample for certain populations.

5. **I want to know more about the new tobacco products (e-cigarettes and heat-not-burn products).**
Questions about e-cigarette use have been included in the TUS-CPS since 2011. The 2018-2019 TUS-CPS included new questions categorizing e-cigarette device type; policies about e-cigarette use in the workplace and rules inside the home; and exposure to federal court-mandated tobacco company statements about the dangers of smoking cigarettes. In the May 2019 questionnaire, two questions about heat-not-burn products were added. Some recent data about new tobacco products are highlighted in the 2014-2015 TUS-CPS Highlights Document, the July 2018 Data Brief, and the 2018-2019 Data Brief. Questionnaires, data files, and brief data reports for all TUS-CPS cycles are available at https://cancercontrol.cancer.gov/brp/tcrb/tus-cps/questionnaires.html.

6. **Would questions on marijuana use be included in future TUS surveys because of the common co-use of marijuana and cigarettes/e-cigarettes? If the answer is no, what data sources which contain marijuana use could be linked with TUS data?**
Co-use of marijuana with cigarettes/e-cigarettes and other tobacco products is a topic of great interest in the research community and may be considered for inclusion in future TUS-CPS cycles. Planning for the 2022-2023 TUS-CPS is currently underway. The CPS does not presently include any surveys with marijuana questions that can be linked with TUS data. Currently, national-level data on marijuana use are captured by the National Survey on Drug Use and Health (NSDUH), the Monitoring the Future (MTF) Study, the Population Assessment of Tobacco and Health (PATH) study, and the Behavioral Risk Factor Surveillance System (BRFSS).

7. **I'm interested in examining the impact of flavor bans on tobacco use and want to explore how the TUS-CPS can be used for this topic.**
The TUS-CPS does not currently collect data on flavor bans. However, this is an important policy evaluation topic that may be considered for the next survey cycle. Beginning with the 2014-2015 TUS-CPS, questions were added to gather data on the consumption of flavored tobacco products. For example, "When you [smoke a cigar/smoke a regular pipe filled with tobacco/smoke a water/hookah pipe filled with tobacco/use an e-cigarette/use smokeless tobacco] is it usually flavored?" Additional questions included in the survey ask respondents to specify the type of flavor consumed (tobacco, menthol, fruit, etc.). For more details, please view the 2014-2015 TUS-CPS questionnaire and more recent questionnaires and data available at https://cancercontrol.cancer.gov/brp/tcrb/tus-cps/questionnaires.html.

8. **What sort of flavors and product design variables are available?**
Since the 2014-2015 cycle, the TUS-CPS has assessed the use of flavored tobacco products. Survey respondents are asked to report whether they consume flavored tobacco products, and which specific flavors they consume. Flavor choices include a) tobacco flavored, b) menthol or mint flavored, c) flavored like clove, spice, herb, fruit, alcohol, candy, sweets, or chocolate, and d) some other type of flavor (if yes, specify).

Questions about product design include asking respondents whether they purchase tobacco products individually (single piece), by the box, or by the carton. Questions about e-cigarette device type were added in the 2018-2019 TUS-CPS, and assess whether consumers use disposable or unknown e-cigarettes, e-cigarette cartridges, or bottles of e-liquid for tank and mod systems.

Additional information and access to TUS-CPS questionnaires and data are available at https://cancercontrol.cancer.gov/brp/tcrb/tus-cps/questionnaires.html.

9. **Is there any way to look at parents and their children in the TUS-CPS?**
   With respect to collection of information on tobacco use by adolescents: From 1992-2006, questions about tobacco use were generally asked of respondents 15 years and older (an exception was January 1996, which interviewed those 16+). Beginning in 2007, the survey was changed to request tobacco use information from adults 18+ only.

   Concerning analysis of family/household dynamics related to tobacco use: A unique feature of the CPS is the collection of household characteristics, including the roster of household members, relationship to head of household for all household members including children 0-14 years of age, and occupation for all respondents 15 years and older. These data may be used to characterize individual TUS respondent records to conduct a variety of analyses on topics such as home smoking and vaping rules, children's exposure to secondhand smoking at home, and more. Additional examples of household analyses are available in the TUS-CPS Publications Database.

Linkages Between TUS-CPS and Other Data

1. **What are best practices for linking TUS-CPS data to other datasets?**
   Because of the CPS's panel design, a subset of persons who were in a sample for any given month of TUS-CPS fielding (e.g., July 2014, January 2015, or May 2015) can be linked with other CPS Basic and Supplement data from surveys conducted within + /-15 months of a TUS-CPS. While there is no step-by-step guide to linking TUS-CPS to other datasets (and the process will vary depending on the other supplement of interest), there are several points to consider when linking your data, shown below.

   General tips:
   - When linking data from two supplements, you will first need to create a common variable for matching observations across datasets. Some good components to include in a unique identifier: month of survey, year of survey (be careful here—the month that an individual takes the TUS and the month they take another supplement will be different, so you may need to edit one or more variables to make month and year match up), household identifiers (HRHHID and HRHHID2), and an individual's line number [PULINENO]). After you've linked the two supplements, check to see that individuals were matched properly by creating a quick set of indicator variables for sex, race, and age.
   - Be careful to check before linking supplements: are there any variables that appear in both datasets that have the same name but represent different things? For example, HRMONTH is generally the variable name for when an individual completes a survey. In the TUS dataset, HRMONTH would indicate when someone participated in TUS. But, because of the panel design of CPS, the same

individual may have participated in the Food Security Supplement (FSS) during a different month. In the FSS dataset, HRMONTH would indicate the month that the participant took the FSS. Linking the two datasets might lead to some confusion, so it can be helpful to plan which variables to include in your linked dataset ahead of time and revise their names as necessary.

On adjusting weights:
- Each month of TUS-CPS data is weighted to the U.S. population. This means that when combining months—whether combining the three months in a wave or combining whichever months make sense for your linkage/analytic goals—the person weight and the replicate weights need to be divided by the number of months being combined. So, if you're combining the 2010-11 (May '10, Aug '10, Jan '11) and 2014-15 (July '14, Jan '15, May '15) waves, the person weight and replicate weights all need to be divided by 6. This can be further complicated when combining waves of the TUS that have different numbers of replicate weights. Guidance on how to do this, including sample code, can be found on the TUS-CPS website here: https://surveillance.cancer.gov/reports/tech2020.01.pdf.
- Depending on how complete linkage is—in other words, what proportion of respondents have their TUS data successfully matched with their responses to other supplements—it may be necessary to account for loss of participants by adjusting the weights further. The TUS team has compiled a helpful resource here: https://cancercontrol.cancer.gov/brp/tcrb/tus-cps/TUS-CPS_overlap.pdf. Though this document deals specifically with panel overlap, the same approach can be applied to adjusting weights after linkage as well.

By no means is this an exhaustive guide to linking datasets, but the documentation on the TUS-CPS and CPS websites can serve as a great resource, and the TUS Team is always just an email away. Additional information about linkages between TUS-CPS and other data sources is available at https://cancercontrol.cancer.gov/brp/tcrb/tus-cps/linkages.html.

2. **What types of summary data are available with geocodes (especially to census tract level), if any?**
   The CPS data items are contained on the TUS data files. These items are already available or can be linked to other CPS files. These geographic items include information on state, metropolitan status (metropolitan, non-metropolitan, not identified), FIPS codes/counties, and some specific metropolitan identifiers (cities or groups of cities, no census tracts). It is worth noting that in the publicly available data, metropolitan status and FIPS code remain unidentified for many observations to protect participants' privacy. Congressional law prohibits the Census Bureau from disclosing geographical information on anyone living in a geographic area of less than 100,000 population. The result is that most county codes are not identified. Geography data is provided to the state level and some sub-state levels for specific metropolitan identifiers (see attachment 11 of the technical documentation: https://www2.census.gov/programs-surveys/cps/techdocs/cpsmay15.pdf). The lowest level with full information is the state level. Please feel free to reach out with additional questions to ncidccpsbrpadvances@mail.nih.gov.

3. **I would like to know your thoughts about the comparability of the questionnaire for some outcomes.**
   TUS-CPS data may be compared to or linked with other CPS supplements and several datasets that provide health outcomes. Details about linking TUS-CPS data with other data sources are available at https://cancercontrol.cancer.gov/brp/tcrb/tus-cps/linkages.html.

4. **What proportion of TUS respondents can be linked to the Annual Social and Economic Supplement (ASEC) supplement?**
Due to the unique structure of CPS, the proportion of individuals whose TUS responses can be linked with responses to other supplements will vary depending on the TUS wave and the other supplement of interest. Using the March 2015 ASEC as an example, only TUS respondents from January and May 2015 can be matched to ASEC due to the timing of each supplement and the 4-8-4 design of CPS (four months on, eight off, four on). For the January and March 2015 TUS, though, there is a 91.4% match for self-respondents using the unweighted data and a 90.8% match for self-respondents using the weighted data.

5. **Are there any pregnancy or parent datasets the TUS-CPS can link to?**
TUS-CPS can link to any other CPS supplements (for complete list with brief descriptions, visit https://www.census.gov/programs-surveys/cps/about/supplemental-surveys.html). CPS supplements most relevant to pregnancy and parenthood would be the fertility supplement (https://catalog.data.gov/dataset/current-population-survey-fertility-supplement) and the Annual Social and Economic Supplement (ASEC, https://www.census.gov/data/datasets/time-series/demo/cps/cps-asec.html), though these may not deal with pregnancy specifically.

Data Harmonization Using TUS-CPS

1. **How can multiple TUS-CPS files be combined and analyzed over several years to better understand trends in adult tobacco product use?**
Data from all currently available waves (1992-2015) were harmonized into a single dataset. Harmonized topics include cigarette use, workplace and home smoking restrictions, attitudes toward smoke-free policies, advice to quit by physician/dentist, health perceptions/beliefs, and other tobacco product use. The harmonized dataset, variable codebook, and associated technical guidance may be found at https://cancercontrol.cancer.gov/brp/tcrb/tus-cps/questionnaires.html. Later survey waves will be added to the harmonized file as they become available.

2. **What are the challenges for data harmonization?**
There are several main challenges. First, as items used to measure the same construct could have changed over time, determining whether and how to reconcile the differences between items and responses can be challenging. In the process, the team evaluates the consistency of items and responses over time, and the interpretability of the harmonized items and responses. Second, the universe of respondents changes between waves, so the team must determine whether and how to construct a meaningful universe of respondents for harmonization. Third, some variables are only assessed in a single wave and therefore not included in the harmonized data. Finally, some harmonized variables may not be useful to researchers for a specific item (e.g., cigarette brand).

3. **I am interested in knowing how to evaluate trends, especially because some questions change over time.**
Details about how questions with different wording were harmonized may be found in the harmonization technical guidance and variable crosswalk codebook, available at https://cancercontrol.cancer.gov/brp/tcrb/tus-cps/questionnaires.html.

4. **From the researchers' perspective, are there any downsides to using the harmonized version of the TUS available through IPUMS?**
   To our knowledge, TUS data are not harmonized in IPUMS. For further information, please consult a contact at IPUMS (https://ipums.org/). Data files and the harmonized dataset are available on the National Cancer Institute's TUS-CPS webpage at https://cancercontrol.cancer.gov/brp/tcrb/tus-cps/questionnaires.html.

Using SAS to Analyze TUS-CPS Data

1. **I'd love to talk about calculating small area estimates for counties using the state-level data.**
   Guidance on calculating small area estimates is available in a PowerPoint presentation from 2018, available at https://cancercontrol.cancer.gov/brp/tcrb/tus-cps/Liu-TUSCPS-TCRB2018.pptx, and an updated presentation from the 2019 SRNT Workshop, available at https://cancercontrol.cancer.gov/brp/tcrb/tus-cps/workshops/2019/TUS-SRNT2019WorkshopPresent.pdf. You may submit more specific questions to ncidccpsbrpadvances@mail.nih.gov.

2. **Do you have any guidance for using sample weights properly in an analysis?**
   Guidance on conducting analyses using sample weights can be found in the 2009 Users' Workshop Presentations, available at https://cancercontrol.cancer.gov/brp/tcrb/tus-cps/workshops/2009/presentations.html. Please refer to the second presentation listed, "Tips and Tricks of Handling the TUS Data." Accompanying SAS code and examples are also provided.

3. **How should we deal with the complex sampling design when pooling the 2018 data with previous TUS data? Would you provide an example in SAS?**
   Data from all currently available waves (1992-2015) were harmonized into a single dataset, available at https://cancercontrol.cancer.gov/brp/tcrb/tus-cps/questionnaires.html. 2018-2019 data will be added shortly. Additional materials and guidance are featured in the 2009 User Workshop (https://cancercontrol.cancer.gov/brp/tcrb/tus-cps/workshops/2009/presentations.html) and the 2013 TUS-CPS Webinar (http://www.cancermeetings.org/TUSCPSWebinar/presentations.cfm). Data files (.dat) and SAS code for each wave of the TUS-CPS are available at https://cancercontrol.cancer.gov/brp/tcrb/tus-cps/questionnaires.html.

4. **I would like to understand better how the response weights/survey weighting in the TUS-CPS works, and how to infer unweighted population-size estimates for adult tobacco product use from these data.**
   Please see the TUS-CPS Technical Documentation for individual survey waves, also found on the Census FTP site—Appendix 16: Source and Accuracy Statement. Also, the webinar linked here should be helpful in describing how the replicate weights were derived: http://www.cancermeetings.org/TUSCPSWebinar/documents/tuscps_Webinar_Liu.pdf.