Sexuality, Health, and Opportunity in the Life Course
Add Health

Sexual Orientation/ Gender Identity, Socioeconomic Status, and Health across the Life Course (SOGI-SES) User Guide: Study Overview

Report prepared by

Carolyn T. Halpern

Dana R. Burshell

Stephanie M. Hernandez

Lixin Qu

Kerith J. Conron

UNC CAROLINA POPULATION CENTER

UCLA School of Law Williams Institute

Add Health
The National Longitudinal Study of Adolescent to Adult Health

CAROLINA POPULATION CENTER | CAROLINA SQUARE - SUITE 210 | 123 WEST FRANKLIN STREET | CHAPEL HILL, NC 27516

Add Health is supported by grant P01-HD31921 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development, with cooperative funding from 23 other federal agencies and foundations.

https://doi.org/10.17615/3971-vk43

Suggested Citations

User Guide: Halpern CT, Burshell DR, Hernandez SM, Qu L, Conron KJ. 2023. Sexual Orientation, Gender Identity, Socioeconomic Status, and Health across the Life Course Study (SOGI-SES) User Guide. The Carolina Population Center, University of North Carolina at Chapel Hill, NC and the Williams Institute, UCLA, Los Angeles, CA funded by the National Institutes of Health. Available from: https://doi.org/10.17615/3971-vk43.

Survey: Conron KJ, Burshell DR, Halpern CT. 2023. Sexual Orientation, Gender Identity, Socioeconomic Status, and Health across the Life Course Study (SOGI-SES) Survey. The Carolina Population Center, University of North Carolina at Chapel Hill, NC and the Williams Institute, UCLA, Los Angeles, CA funded by the National Institutes of Health. Available from: https://doi.org/10.17615/6dtv-qg09.

Contents

Suggested Citations 2

Contents 2

Introduction 3

Study Design 4

Data Collection 8

Final Data Set Inclusion Criteria 9

Weighting and Non-Bias Analysis 10

Dissemination of Study Materials 12

Using the Codebooks 13

Analyzing the Data 18

Resources 20

Acknowledgements 20

The Sexual Orientation/Gender Identity, Socioeconomic Status, and Health Across the Life Course (SOGI-SES) Study is co-directed by Principal Investigators Carolyn T. Halpern at the University of North Carolina at Chapel Hill and Kerith J. Conron at the Williams Institute, UCLA and is funded by the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (NICHD) and the National Institute on Minority Health and Health Disparities under grants R01 HD087365 and R01 HD087365-03S1. SOGI-SES is an Add Health ancillary study that includes a subset of Add Health respondents.

Waves I-V of Add Health were funded by grant P01 HD31921 (Harris) from the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (NICHD), with cooperative funding from 23 other federal agencies and foundations. Add Health is currently directed by Robert A. Hummer and funded by the National Institute on Aging cooperative agreements U01 AG071448 (Hummer) and U01AG071450 (Aiello and Hummer) at the University of North Carolina at Chapel Hill. Add Health was designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris at the University of North Carolina at Chapel Hill.

Introduction

The Sexual Orientation/Gender Identity, Socioeconomic Status, and Health across the Life Course Study (SOGI-SES) collected new survey data to support exploration of the relationships among sexual orientation, gender identity, gender expression, same-sex romantic and sexual behaviors, socioeconomic status, and health. The study was designed and conducted under the direction of the co-PIs, Carolyn T. Halpern from the Carolina Population Center (CPC) at the University of North Carolina at Chapel Hill and Kerith J. Conron from the Williams Institute at the University of California Los Angeles (UCLA). SOGI-SES is funded by a grant from the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development and the National Institute on Minority Health and Health Disparities (Grant numbers: 1R01HD087365 and R01 HD087365-03S1).

Based on a subset of respondents in the National Longitudinal Study of Adolescent to Adult Health (Add Health) cohort, the purposes of SOGI-SES were to: 1) develop a new survey focused on formative developmental experiences related to sexual orientation and gender identity, and determinants of socioeconomic status (SES); 2) collect, clean, document, and disseminate these new data; 3) conduct analyses to explore SES, stigma, and discrimination as contributors to mental health disparities among sexual and gender minorities.

Housed at the Carolina Population Center, Add Health is an innovative longitudinal study that has followed a large population-based sample of US adolescents well into adulthood to collect information about determinants of health and well-being. In addition to gathering information about health behaviors, Add Health captured information about the social and physical contexts in which individuals live and grow. The study was a leader in early efforts to collect biological specimens from large, population-based samples. Add Health used a school-based design that selected a random sample of US schools and administered surveys in 1994-1995 to more than 90,000 students in 7th through 12th grades. From the school rosters, a random sample of adolescents and one parent was selected for in-home interviews; specific subpopulations were oversampled for research on vulnerable and rare populations, yielding an “in-home” interview sample of almost 21,000 students who have since been followed with five completed surveys to date. At this writing, a sixth survey is underway. Table 1 provides an overview of Add Health and SOGI-SES data collection. SOGI-SES used Add Health Wave V participants as its sampling frame. The Wave V survey was conducted between 2016 and 2018. For more information about Add Health such as study design, user guides, codebooks, and data, please visit https://addhealth.cpc.unc.edu/.

The SOGI-SES study was designed to collect data that, paired with Add Health’s extensive prospective longitudinal data, will inform public health policy and intervention strategies to reduce the prevalence of disparities in health for sexual and gender minority populations. The design, planning, and implementation of the study were performed through a collaboration between the Carolina Population Center (CPC), the Williams Institute, and RTI International (RTI; field contractor). This multilevel, longitudinal intergenerational data resource provides unprecedented research opportunities to learn about sexual and gender minority health and development across the life course.

UNC, UCLA, and RTI received Institutional Review Board (IRB) approval for the study. UCLA and RTI relied on UNC’s IRB, the lead IRB for this study. All procedures, survey content, and informed consent forms

The Sexual Orientation/Gender Identity, Socioeconomic Status, and Health Across the Life Course (SOGI-SES) Study is co-directed by Principal Investigators Carolyn T. Halpern at the University of North Carolina at Chapel Hill and Kerith J. Conron at the Williams Institute, UCLA and is funded by the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (NICHD) and the National Institute on Minority Health and Health Disparities under grants R01 HD087365 and R01 HD087365-03S1. SOGI-SES is an Add Health ancillary study that includes a subset of Add Health respondents.

Waves I-V of Add Health were funded by grant P01 HD31921 (Harris) from the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (NICHD), with cooperative funding from 23 other federal agencies and foundations. Add Health is currently directed by Robert A. Hummer and funded by the National Institute on Aging cooperative agreements U01 AG071448 (Hummer) and U01AG071450 (Aiello and Hummer) at the University of North Carolina at Chapel Hill. Add Health was designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris at the University of North Carolina at Chapel Hill.

were reviewed and approved by the UNC IRB. Data safety and confidentiality were ensured by following Add Health's Data Security Plan, which can be found on the Add Health website.

Table 1: Overview of Add Health and SOGI-SES Data Collection Efforts*

Year	Age of Respondents	Data Collection Efforts: Add Health Waves and SOGI-SES*
1994	11–18	Add Health Wave I—In-school interviews of adolescents in grades 7 through 12.
1995	12–19	Add Health Wave I—In-home interviews of adolescents and parents.
1996	13–20	Add Health Wave II—In-home follow-up interviews with adolescents.
2001–2002	18–26	Add Health Wave III—In-home interviews with respondents from Wave I; subset of romantic partners. Biospecimens and GPS location collected.
2007–2008	24–32	Add Health Wave IV—In-home interviews with respondents from Wave I. Biospecimens and GPS location collected.
2015–2016	50–80	Parent Study- In-home interviews of a sample of the Add Health parents interviewed at Wave I. Cohabiting spouse/partners interviews also conducted.
2016–2018	31–44	Add Health Wave V—Multimode data collection (web, mail, in person, and phone) with Wave I respondents. Biospecimens, GPS location, and consent to obtain birth records for those born in 6 states collected.
2020–2021	35–45	SOGI-SES – Web-only survey with a subset of Add Health Wave V participants.
2022–2023	~ 37–48**	Add Health Wave VI - Multimode data collection (web, in person) with Wave I respondents. Biospecimens and GPS location collected as well as consent to obtain birth and death records.

* Other Add Health ancillary studies, not shown in Table 1, have been conducted.

** Predicted age

Study Design

Sampling Frame

The SOGI-SES eligible sampling frame consisted of living Add Health Wave V respondents who have Wave I sampling weights (n=11,712).

Sexual and Gender Minority (SGM) Definition and Inclusion

Figure 1 shows the SOGI-SES definitions and sampling strategy.

For the purposes of the study, sexual and gender minorities (SGM) were defined as respondents meeting any of the following criteria:

- Sexual identity of mostly heterosexual, bisexual, mostly homosexual, or 100% homosexual at Add Health Waves III, IV, and/or V

The Sexual Orientation/Gender Identity, Socioeconomic Status, and Health Across the Life Course (SOGI-SES) Study is co-directed by Principal Investigators Carolyn T. Halpern at the University of North Carolina at Chapel Hill and Kerith J. Conron at the Williams Institute, UCLA and is funded by the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (NICHD) and the National Institute on Minority Health and Health Disparities under grants R01 HD087365 and R01 HD087365-03S1. SOGI-SES is an Add Health ancillary study that includes a subset of Add Health respondents.

Waves I-V of Add Health were funded by grant P01 HD31921 (Harris) from the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (NICHD), with cooperative funding from 23 other federal agencies and foundations. Add Health is currently directed by Robert A. Hummer and funded by the National Institute on Aging cooperative agreements U01 AG071448 (Hummer) and U01AG071450 (Aiello and Hummer) at the University of North Carolina at Chapel Hill. Add Health was designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris at the University of North Carolina at Chapel Hill.

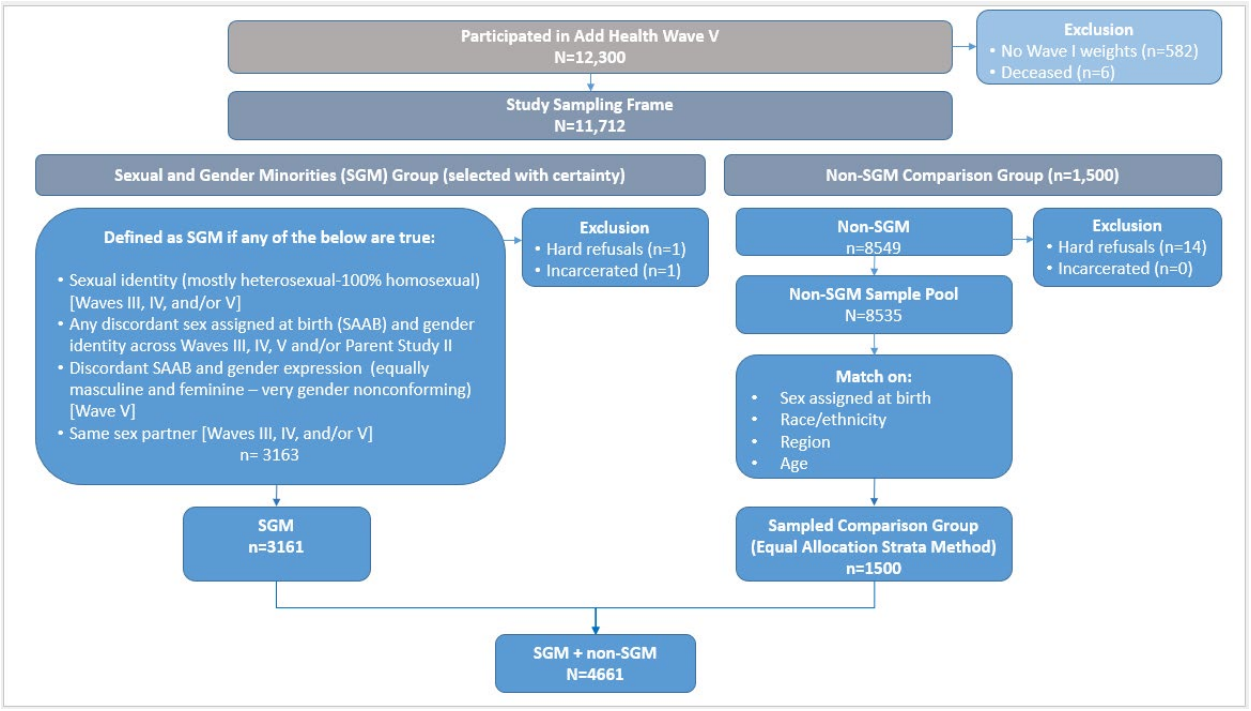
- Any discordant sex assigned at birth and gender identity across Waves III, IV, V and/or the Add Health Parent Study II
- Same sex partner (Waves III, IV, and/or V)
- Discordant sex assigned at birth and gender expression (equally masculine and feminine – very gender nonconforming) (Wave V) is further defined in Table 2.

Table 2. Defining Concordant Sex Assigned at Birth and Gender Expression

Gender Expression Response Options	Concordance for Sex Assigned at Birth Females	Concordance for Sex Assigned at Birth Males
Very feminine	Very gender conforming	Very gender nonconforming
Mostly feminine	Mostly conforming	Mostly nonconforming
Somewhat feminine	Somewhat conforming	Somewhat nonconforming
Equally feminine and masculine	Equally feminine and masculine	Equally feminine and masculine
Somewhat masculine	Somewhat nonconforming	Somewhat conforming
Mostly masculine	Mostly nonconforming	Mostly conforming
Very masculine	Very gender nonconforming	Very gender conforming

Wave V respondents who met the SGM study definition and were neither incarcerated nor had indicated a hard study participation refusal, were selected with certainty (n=3,161). The remaining non-SGMs (n=8,549) were sampled to create a comparison sample of 1,500. In total, there were 4,661 Wave V respondents selected to participate in the SOGI-SES survey.

Figure 1. Study Inclusion, Exclusion, and Sampling Algorithm



The Sexual Orientation/Gender Identity, Socioeconomic Status, and Health Across the Life Course (SOGI-SES) Study is co-directed by Principal Investigators Carolyn T. Halpern at the University of North Carolina at Chapel Hill and Kerith J. Conron at the Williams Institute, UCLA and is funded by the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (NICHD) and the National Institute on Minority Health and Health Disparities under grants R01 HD087365 and R01 HD087365-03S1. SOGI-SES is an Add Health ancillary study that includes a subset of Add Health respondents.

Waves I-V of Add Health were funded by grant P01 HD31921 (Harris) from the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (NICHD), with cooperative funding from 23 other federal agencies and foundations. Add Health is currently directed by Robert A. Hummer and funded by the National Institute on Aging cooperative agreements U01 AG071448 (Hummer) and U01AG071450 (Aiello and Hummer) at the University of North Carolina at Chapel Hill. Add Health was designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris at the University of North Carolina at Chapel Hill.

Sampling of non-SGM

The Equal Allocation approach was used to sample non-SGM because it provided the best compromise of minimizing overall unequal weighting effects (UWEs) and minimizing the model prediction error for subsequent analyses. The non-SGM sampling frame was stratified by race/ethnicity, sex assigned at birth and poverty level to form 16 strata. These variables were chosen to optimize representation based on sex assigned at birth and race. In addition, because a project aim was to enhance exploration of socioeconomic contributors to health, it was important to oversample those with a lower socioeconomic status, defined as under 200% of the federal poverty level.

To select the non-SGM sample, the 8,535 sample members were placed into the 16 strata shown in Table 3, and, using probability proportional to size sampling in each stratum, cases were selected in each stratum as shown in the last column of Table 3. The size measure used for this proportional size sample was the Wave V Grand Sample Weight (GSWGT5).

Table 3. Final Sample Sizes by Stratum for Equal Allocation Method

Stratum	Sex assigned at birth	Race/Ethnicity	Poverty	Non-SGM Frequency	Non-SGM Sample Size
111	Male	White	<200 FPL	445	97
112	Male	White	>=200 FPL	2,029	97
121	Male	Black	<200 FPL	265	96
122	Male	Black	>=200 FPL	471	96
131	Male	Hispanic	<200 FPL	134	96
132	Male	Hispanic	>=200 FPL	477	97
141	Male	Other	<200 FPL	79	79
142	Male	Other	>=200 FPL	365	96
211	Female	White	<200 FPL	506	97
212	Female	White	>=200 FPL	1,766	97
221	Female	Black	<200 FPL	474	97
222	Female	Black	>=200 FPL	530	97
231	Female	Hispanic	<200 FPL	215	96
232	Female	Hispanic	>=200 FPL	443	97
241	Female	Other	<200 FPL	69	69
242	Female	Other	>=200 FPL	267	96
Total				8,535	1,500

Survey Development

In 2019, RTI conducted 27 one-on-one interviews to assess the content, understandability, and usability of the web-based SOGI-SES survey instrument. Participants were recruited from the Research Triangle in

The Sexual Orientation/Gender Identity, Socioeconomic Status, and Health Across the Life Course (SOGI-SES) Study is co-directed by Principal Investigators Carolyn T. Halpern at the University of North Carolina at Chapel Hill and Kerith J. Conron at the Williams Institute, UCLA and is funded by the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (NICHD) and the National Institute on Minority Health and Health Disparities under grants R01 HD087365 and R01 HD087365-03S1. SOGI-SES is an Add Health ancillary study that includes a subset of Add Health respondents.

Waves I-V of Add Health were funded by grant P01 HD31921 (Harris) from the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (NICHD), with cooperative funding from 23 other federal agencies and foundations. Add Health is currently directed by Robert A. Hummer and funded by the National Institute on Aging cooperative agreements U01 AG071448 (Hummer) and U01AG071450 (Aiello and Hummer) at the University of North Carolina at Chapel Hill. Add Health was designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris at the University of North Carolina at Chapel Hill.

North Carolina and Chicago, IL, could not be Add Health participants, and were selected to include diversity in gender (cisgender, transgender), sexual orientation (straight, lesbian/gay/bisexual (LGB)), race/ethnicity (especially non-Hispanic white, non-Hispanic Black, Hispanic), and education level.

Interview feedback was used to revise the survey instrument. Overall, feedback confirmed that people across differing sexes and genders (i.e., cisgender, straight, SGM) felt comfortable understanding the questions, answering the questions, and using the web survey.

Survey Mode and Content

SOGI-SES used an online survey; there was no paper survey option. The programmed survey uses extensive skip patterns to simplify content flow and decrease burden for respondents. All survey questions, skip logic, and programmer notes are included below. For data analysis purposes, we recommend using this User Guide document in combination with the study's Codebooks, Index, and associated Survey available at Add Health's website - https://addhealth.cpc.unc.edu/.

Table 4 displays survey sections and organization. Note that some sections and variables do not have a number and are not included in the Codebooks. Questions given the QNA (Question Not Available) designation are not included in the Codebooks due to confidentiality issues, or because they are for internal use only. Examples include, but are not limited to, questions related to identity confirmation, geography, write-ins, and contact information.

The Sexual Orientation/Gender Identity, Socioeconomic Status, and Health Across the Life Course (SOGI-SES) Study is co-directed by Principal Investigators Carolyn T. Halpern at the University of North Carolina at Chapel Hill and Kerith J. Conron at the Williams Institute, UCLA and is funded by the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (NICHD) and the National Institute on Minority Health and Health Disparities under grants R01 HD087365 and R01 HD087365-03S1. SOGI-SES is an Add Health ancillary study that includes a subset of Add Health respondents.

Waves I-V of Add Health were funded by grant P01 HD31921 (Harris) from the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (NICHD), with cooperative funding from 23 other federal agencies and foundations. Add Health is currently directed by Robert A. Hummer and funded by the National Institute on Aging cooperative agreements U01 AG071448 (Hummer) and U01AG071450 (Aiello and Hummer) at the University of North Carolina at Chapel Hill. Add Health was designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris at the University of North Carolina at Chapel Hill.

Table 4: Survey Sections

Number	Name	Letter Designation
NA*	Identity Verification	NA*
1	Education	ED
2	COVID-19	CO
3	Employment	EM
4	Gender Identity	GI
5	Sexual Orientation	SO
6	Gender Expression	GE
7	Household Roster	HR
8	Non-Household Child Roster	CR
9	Children's Healthcare Needs	CH
10	Romantic Partner	P
11	Mental Health	MH
12	Perceived Discrimination	PD
13	Economics	EC
14	Intergenerational Transfers	IG
15	Sexual Behavior	SB
16	Parental Acceptance	PA
17	Minority Stress	MS
18	Family Formation	FF
19	Adverse Experience	AE
20	Health	HE
NA*	Contact Information	NA*

* NA – Not applicable, questions are not in Codebook due to confidentiality issues.

Rosters

Section 7: *Household Roster* and Section 8: *Non-Household Child Roster* can each accommodate up to 12 unique persons. Section 9: *Children's Healthcare Needs* pulls data from the children in the Household and/or Non-Household Child rosters for applicable questions. Note that *Children's Healthcare Needs* individuals 1-12 correspond to the 1-12 positions in the *Household Roster*. *Children's Healthcare Needs* individuals 13-24 correspond to the 1-12 positions in the *Non-Household Child Roster*. For example, Child #6 in the *Children's Healthcare Needs* refers to Member #6 in the *Household Roster*. Child #16 in the *Children's Healthcare Needs* refers to the NH Child #4 in the *Non-Household Child Roster*.

Data Collection

Tracing

RTI conducted tracing before and during "field work" (i.e., period when the survey was open) to achieve the highest level of accurate contact information. Eligible participants were notified about the web survey either by email (88.2%) or mail (if they did not participate in Wave V via web or if an email was not

The Sexual Orientation/Gender Identity, Socioeconomic Status, and Health Across the Life Course (SOGI-SES) Study is co-directed by Principal Investigators Carolyn T. Halpern at the University of North Carolina at Chapel Hill and Kerith J. Conron at the Williams Institute, UCLA and is funded by the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (NICHD) and the National Institute on Minority Health and Health Disparities under grants R01 HD087365 and R01 HD087365-03S1. SOGI-SES is an Add Health ancillary study that includes a subset of Add Health respondents.

Waves I-V of Add Health were funded by grant P01 HD31921 (Harris) from the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (NICHD), with cooperative funding from 23 other federal agencies and foundations. Add Health is currently directed by Robert A. Hummer and funded by the National Institute on Aging cooperative agreements U01 AG071448 (Hummer) and U01AG071450 (Aiello and Hummer) at the University of North Carolina at Chapel Hill. Add Health was designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris at the University of North Carolina at Chapel Hill.

available). If the email or mailing attempt led to the communication being returned, additional tracing efforts were conducted.

Recruitment and Timeline

Recruitment efforts spanned between August 25, 2020 and February 24, 2021. The online survey was open to respondents September 1, 2020 through February 28, 2021. During recruitment, a combination of emails and various mailed items (letters, folded postcards, and a "FedEx"-like envelope) were sent as reminders to start or complete the survey. Ultimately, participants could have been contacted about the study up to 13 times (1 pre-notification email/letter, 1 recruitment email/letter, 7 email reminders and 4 mailed items).

COVID-19

Data collection occurred during the first year of the COVID-19 pandemic. COVID-19 isolation and safety precautions significantly changed how people worked, socialized, and cared for their loved ones, likely placing a higher burden on families with children. It is possible that these changes impacted survey participation.

Individual Passcode and Incentive

Add Health respondents eligible for the study were provided with an individual passcode to access the survey and could stop and start the survey as needed. Respondents who completed the survey received a $50 online VISA gift card.

Final Data Set Inclusion Criteria

Of the 12,300 Add Health Wave V participants, 4,661 were selected to participate in the SOGI-SES survey. Of those, nonrespondents included those who contacted the study to refuse participation, did not open the survey, or did not provide informed consent after opening the survey. Respondents were also excluded from the data set if they failed identity confirmation.

A survey was defined as "complete," and therefore to be included in the final released data set, if a respondent completed the survey through the Mental Health survey section (Section 11 of 20; see section on survey development). The Mental Health section was essential to the project's primary interest in mental health outcomes. Of the 116 survey breakoffs, 77 occurred after the Mental Health section. These 77 surveys were therefore defined as complete and included in the final data set.

Table 5 provides additional information about the final disposition status of Wave V Add Health respondents selected for SOGI-SES participation.

The Sexual Orientation/Gender Identity, Socioeconomic Status, and Health Across the Life Course (SOGI-SES) Study is co-directed by Principal Investigators Carolyn T. Halpern at the University of North Carolina at Chapel Hill and Kerith J. Conron at the Williams Institute, UCLA and is funded by the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (NICHD) and the National Institute on Minority Health and Health Disparities under grants R01 HD087365 and R01 HD087365-03S1. SOGI-SES is an Add Health ancillary study that includes a subset of Add Health respondents.

Waves I-V of Add Health were funded by grant P01 HD31921 (Harris) from the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (NICHD), with cooperative funding from 23 other federal agencies and foundations. Add Health is currently directed by Robert A. Hummer and funded by the National Institute on Aging cooperative agreements U01 AG071448 (Hummer) and U01AG071450 (Aiello and Hummer) at the University of North Carolina at Chapel Hill. Add Health was designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris at the University of North Carolina at Chapel Hill.

Table 5: Final Disposition of Selected Survey Participants (n=4661)

Final Disposition	Frequency
Respondents Included in Final Data Set	**2614**
Completed Survey	2537
Partial, but met definition for complete (Completed up to Section 12)	77
Respondents Excluded from Final Data Set	**87**
Failed Identity Confirmation	48
Partials (Did not complete up to Section 12)	39
Nonrespondents	**1929**
Did Not Open Survey	1873
Did Not Consent	48
Refusals	8
Ineligible	**31**
Deceased	29
Incarcerated	2
Total	**4661**

Response Rate

Our conservative response rate, calculated as those included in the final data set divided by the total number of eligible respondents, was 56.5% (2614/4630=56.5%).

Weighting and Non-Bias Analysis

RTI International developed cross-sectional survey weights for the SOGI-SES sample and conducted a non-response bias analysis. The relative bias was deemed "moderately small," and the validation analysis confirmed that weighted estimates using final analysis weights were relatively accurate and reliable. Due to budget constraints, longitudinal weights were not developed. Below is a summary of the weighting and non-bias analysis process. For a detailed description see the Guidelines for Analyzing Add Health Data (doi.org/10.17615/C6BW8W) available on the Add Health website.

Sampling Frame for Weighting

The SOGI-SES sampling frame contained all the Add Health Wave V sample members who had completed the Wave V interview with some exclusions. Two different sampling frames were constructed, one for sampling, as described earlier, and one for weighting. This is because some Wave V respondents are considered eligible (i.e., within the scope) for weighting, but not eligible for sampling. As seen in Table 6, these cases are included in the sampling frame for weighting (Eligible to be included in the Frame) but excluded in the sampling frame for sampling (Eligible to be Fielded).

The Sexual Orientation/Gender Identity, Socioeconomic Status, and Health Across the Life Course (SOGI-SES) Study is co-directed by Principal Investigators Carolyn T. Halpern at the University of North Carolina at Chapel Hill and Kerith J. Conron at the Williams Institute, UCLA and is funded by the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (NICHD) and the National Institute on Minority Health and Health Disparities under grants R01 HD087365 and R01 HD087365-03S1. SOGI-SES is an Add Health ancillary study that includes a subset of Add Health respondents.

Waves I-V of Add Health were funded by grant P01 HD31921 (Harris) from the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (NICHD), with cooperative funding from 23 other federal agencies and foundations. Add Health is currently directed by Robert A. Hummer and funded by the National Institute on Aging cooperative agreements U01 AG071448 (Hummer) and U01AG071450 (Aiello and Hummer) at the University of North Carolina at Chapel Hill. Add Health was designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris at the University of North Carolina at Chapel Hill.

Table 6. Number of Cases by Strata and Disposition Code among the Wave V Respondents at the Time when Sampling for the SOGI-SES Study

SGM Group	Disposition	Number of Wave V Respondents	Eligible to be included in the Frame	Eligible to be Fielded
--	Have no Wave I grand weight	582	√	
Non-SGM	Complete W5	8,535	√	√
Non-SGM	Deceased	4		
Non-SGM	Refusal-Hostile	14	√	
SGM	Complete W5	3,161	√	√
SGM	Deceased	2		
SGM	Refusal-Hostile	1	√	
SGM	Subject Incarcerated	1	√	
Total		**12,300**		

Weighting Process

Because the sample was drawn after stratifying by SGM and non-SGM sample members (i.e., SOGI-SES sample members), **the weights were created by SGM and non-SGM groups separately**. The weighting process to create the final analysis weights for the SOGI-SES respondents within each group contains three major steps, including:

1) construction of the base weights to reflect the initial sample design;
2) nonresponse adjustment to adjust for selected cases who did not respond to the survey;
3) final calibration to adjust the weights to meet control totals calculated based on the Add Health Wave V data.

To account for cases that have missing Wave I grand weights and were not assigned to either stratum, the final weights generated from the weighting process described above were multiplied by 1.000568. This value is the ratio of the sum of the Wave V Sample Grand Weights (GSWGT5) of all the cases in the frame (including cases with missing Wave I grand weights) to the sum of the final calibrated weights.

Quality Control and Assessment of the Final Weighting Results

Since the final analysis weights for SOGI-SES survey data were produced through a complex multi-step process, a comprehensive and exhaustive quality control (QC) evaluation was performed to provide ample assurance that all steps of the process were performed as specified and the results were as expected. The final weights as well as all the interim weights created in the process of generating the final weights were checked and verified at the macro- to the micro-levels in the QC procedure.

To validate the impact of weighting on the final estimates of study variables, the weighted estimates and their relative standard errors (RSEs) were compared across the base weights and the final analysis weights by the SGM and non-SGM strata. Virtually all the estimates are very close between the base weights and the final analysis weights with only one appreciable discrepancy. A relative bias, which is the difference

The Sexual Orientation/Gender Identity, Socioeconomic Status, and Health Across the Life Course (SOGI-SES) Study is co-directed by Principal Investigators Carolyn T. Halpern at the University of North Carolina at Chapel Hill and Kerith J. Conron at the Williams Institute, UCLA and is funded by the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (NICHD) and the National Institute on Minority Health and Health Disparities under grants R01 HD087365 and R01 HD087365-03S1. SOGI-SES is an Add Health ancillary study that includes a subset of Add Health respondents.

Waves I-V of Add Health were funded by grant P01 HD31921 (Harris) from the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (NICHD), with cooperative funding from 23 other federal agencies and foundations. Add Health is currently directed by Robert A. Hummer and funded by the National Institute on Aging cooperative agreements U01 AG071448 (Hummer) and U01AG071450 (Aiello and Hummer) at the University of North Carolina at Chapel Hill. Add Health was designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris at the University of North Carolina at Chapel Hill.

between the two sets of the weights divided by the estimates based on the base weights, is usually considered to be small if its absolute value is less than 10%. There is only one variable, H5SE5r (ever anal intercourse), where estimates for the non-SGM stratum exceed this threshold with a relative bias of 13.63%. This relative bias can still be considered as moderately small. In summary, this validation analysis confirms that the weighted estimates with the final analysis weights are relatively accurate and reliable.

Dissemination of Study Materials

User Guide, Study Survey, Codebooks, and Index

SOGI-SES publicly disseminated materials include this User Guide, associated Study Survey, Codebooks, and Index (all publicly available at https://addhealth.cpc.unc.edu/ under Documentation).

Index

The Index is located at the end of the SOGI-SES (Main) Codebook. The Index includes the survey question and the variable name. It is a helpful tool to use in conjunction with the Codebook and Survey.

Data Sets

~~Survey data are only available by contract. There is no public use data set. The cost to be an Add Health contract holder is $1000. Add Health contract holders can access all Add Health and SOGI-SES data approved for use. Please see the Add Health website (https://addhealth.cpc.unc.edu/) and the CPC Data Portal (https://data.cpc.unc.edu) for more information about completing the required paperwork to access Add Health data.~~

Two contractual SOGI-SES data sets were constructed: a **SOGI-SES** data set and a SOGI-SES **Sensitive** data set. They have corresponding SOGI-SES and SOGI-SES Sensitive Codebooks.

SOGI-SES Data Set (Main)

The SOGI-SES data set, which we will refer to as the Main data set in this User Guide, provides the SOGI-SES cross-sectional analysis weight, respondent's age at the time they completed the SOGI-SES survey, the month and year they took the survey, and all survey variables except for sensitive variables that are only available in the Sensitive data set (see below). The sensitive variables are listed in the Main data set and Codebook; however, all the responses were changed to "8888: Request Access." Constructed variables and masked/recoded variables are included in the Main data set and Codebook. Constructed variable names follow section naming conventions; recoded variables are indicated with an "r" at the beginning of the variable name and labeled to indicate a recode. To access the Main data set, users must be Add Health contract holders. Access to the Sensitive data set requires additional Add Health approval.

Sensitive Data Set

The Sensitive data set includes the raw data for variables that were identified as "sensitive" (i.e., those raising deductive disclosure risks related to gender identity, in vitro fertilization, and history of HIV/AIDS

The Sexual Orientation/Gender Identity, Socioeconomic Status, and Health Across the Life Course (SOGI-SES) Study is co-directed by Principal Investigators Carolyn T. Halpern at the University of North Carolina at Chapel Hill and Kerith J. Conron at the Williams Institute, UCLA and is funded by the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (NICHD) and the National Institute on Minority Health and Health Disparities under grants R01 HD087365 and R01 HD087365-03S1. SOGI-SES is an Add Health ancillary study that includes a subset of Add Health respondents.

Waves I-V of Add Health were funded by grant P01 HD31921 (Harris) from the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (NICHD), with cooperative funding from 23 other federal agencies and foundations. Add Health is currently directed by Robert A. Hummer and funded by the National Institute on Aging cooperative agreements U01 AG071448 (Hummer) and U01AG071450 (Aiello and Hummer) at the University of North Carolina at Chapel Hill. Add Health was designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris at the University of North Carolina at Chapel Hill.

diagnosis) and therefore were either removed from the Main data set or masked/recoded to lower disclosure risk.

The Sensitive data set is tightly controlled and access to it requires a special request by Add Health contract holders. Once a request has been approved, users can only work with the raw data and link it to other data while on the secured server. No raw sensitive data can be removed from the secured server. Requests must provide a rationale for access and plan for data use that are reviewed and approved by Add Health. The associated Codebook is publicly available.

Data in the Sensitive data set are available for 1 year to approved users.

Information Not Released

Personal identifiable information was not released. This includes identity confirmation variables, where respondents live, contact information, any write-ins, and initials of household members, children, partners, or parents. Other variables that were used internally such as confirmation questions were also not released.

Using the Codebooks

Codebooks, Index, and Data Set

The Main data set and Codebook follow the Survey Sections (see Table 7) with the addition of Section A: Cross-Sectional Analysis Weight and Survey Information (WS), which includes the SOGI-SES participant's ID, cross-sectional analysis weight, age at time of survey, and the month and year respondent took the survey.

The Sexual Orientation/Gender Identity, Socioeconomic Status, and Health Across the Life Course (SOGI-SES) Study is co-directed by Principal Investigators Carolyn T. Halpern at the University of North Carolina at Chapel Hill and Kerith J. Conron at the Williams Institute, UCLA and is funded by the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (NICHD) and the National Institute on Minority Health and Health Disparities under grants R01 HD087365 and R01 HD087365-03S1. SOGI-SES is an Add Health ancillary study that includes a subset of Add Health respondents.

Waves I-V of Add Health were funded by grant P01 HD31921 (Harris) from the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (NICHD), with cooperative funding from 23 other federal agencies and foundations. Add Health is currently directed by Robert A. Hummer and funded by the National Institute on Aging cooperative agreements U01 AG071448 (Hummer) and U01AG071450 (Aiello and Hummer) at the University of North Carolina at Chapel Hill. Add Health was designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris at the University of North Carolina at Chapel Hill.

Table 7. SOGI-SES (Main) Codebook Sections

Letter/Number	Name	Letter Designation
A	Cross-Sectional Analysis Weight and Survey Information	WS
N/A*	Identity Verification	N/A*
1	Education	ED
2	COVID-19	CO
3	Employment	EM
4	Gender Identity	GI
5	Sexual Orientation	SO
6	Gender Expression	GE
7	Household Roster	HR
8	Non-Household Child Roster	CR
9	Children's Healthcare Needs	CH
10	Romantic Partner	P
11	Mental Health	MH
12	Perceived Discrimination	PD
13	Economics	EC
14	Intergenerational Transfers	IG
15	Sexual Behavior	SB
16	Parental Acceptance	PA
17	Minority Stress	MS
18	Family Formation	FF
19	Adverse Experience	AE
20	Health	HE
N/A*	Contact Information	N/A*

* N/A – Not applicable because not in Codebook due to confidentiality issues.

Coding Conventions

Variable Names:

- **Survey designation** (M1): "M" indicates a SOGI-SES survey variable. The number "1" indicates that it is the first SOGI-SES survey in case there are future surveys.
- **Section** (1-2 capitalized letters): Two capitalized letters correspond to the section content, except for the Romantic Partners section, which only has 1 letter due to the variable name character limits (see Table 7. Codebook Sections).
- **Question** (2-digit number): A 2-digit number indicates specific questions. Additional combinations of letters and numbers may indicate situations like sub-questions and "check all that apply" answer options.

In the Codebook example below (Table 8) for Section 12: Perceived Discrimination (PD), the M1PD06 variable name corresponds to the SOGI-SES survey's (M1) Perceived Discrimination section (PD) question number 6 (06).

The Sexual Orientation/Gender Identity, Socioeconomic Status, and Health Across the Life Course (SOGI-SES) Study is co-directed by Principal Investigators Carolyn T. Halpern at the University of North Carolina at Chapel Hill and Kerith J. Conron at the Williams Institute, UCLA and is funded by the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (NICHD) and the National Institute on Minority Health and Health Disparities under grants R01 HD087365 and R01 HD087365-03S1. SOGI-SES is an Add Health ancillary study that includes a subset of Add Health respondents.

Waves I-V of Add Health were funded by grant P01 HD31921 (Harris) from the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (NICHD), with cooperative funding from 23 other federal agencies and foundations. Add Health is currently directed by Robert A. Hummer and funded by the National Institute on Aging cooperative agreements U01 AG071448 (Hummer) and U01AG071450 (Aiello and Hummer) at the University of North Carolina at Chapel Hill. Add Health was designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris at the University of North Carolina at Chapel Hill.

Table 8: Codebook output for Section 12: Perceived Discrimination Variable M1PD06

M1PD06 - PD06 TIMES BEEN FIRED FROM A JOB- M1 *Interpret as SOGI-SES survey's (M1) Perceived Discrimination section (PD) question number 6 (06).*					
Type		Code			
Measurement Unit		Numeric			
M1PD06		Q6. Since you were 18, how many times have you been fired from a job?			
Logic		if S3Q1≠3, ask Q6 *Interpret as, if the response to Section 3's first question is not equal to 3, ask question 6 in this section.*			
			Frequency	% of total	% of valid
Valid	0	Not applicable, I have never worked for pay	0	0%	0%
	1	Never	1496	57.23%	58.32%
	2	Once	676	25.86%	26.35%
	3	Twice	280	10.71%	10.92%
	4	Three or more times	113	4.32%	4.41%
		Total	2,565	98.13%	100%
Missing	.	Missing *Indicates respondents received question but purposefully skipped. Note: When there are no legitimate skips, Missing could potentially include break offs after Section 11.*	23	0.88%	
	97	Legitimate Skip *Indicates legitimately did not receive the question.*	26	0.99%	
		Total	49	1.87%	
Valid		Invalid	Minimum	Maximum	
2565		49	1	4	

Logic Conventions

If skip logic is limited to within the section, only the question numbers are used. If the skip logic references variables from other sections, the section number and question number is used. To remind yourself what each section number corresponds to, you can use the User Guide's Tables 4 or 7 in this document or refer to the survey or the Main Codebook.

For example, in the M1PD06 example, the logic "if S3Q1≠3, ask Q6" can be interpreted as, if the response to Section 3's question 1 is not equal to 3, ask question 6 in this section. Section 3 is Employment.

Select "all that apply" and Sub-Questions

Questions that specify "select all that apply" and/or have sub-questions only list the general question to simplify logic presentation.

The Sexual Orientation/Gender Identity, Socioeconomic Status, and Health Across the Life Course (SOGI-SES) Study is co-directed by Principal Investigators Carolyn T. Halpern at the University of North Carolina at Chapel Hill and Kerith J. Conron at the Williams Institute, UCLA and is funded by the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (NICHD) and the National Institute on Minority Health and Health Disparities under grants R01 HD087365 and R01 HD087365-03S1. SOGI-SES is an Add Health ancillary study that includes a subset of Add Health respondents.

Waves I-V of Add Health were funded by grant P01 HD31921 (Harris) from the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (NICHD), with cooperative funding from 23 other federal agencies and foundations. Add Health is currently directed by Robert A. Hummer and funded by the National Institute on Aging cooperative agreements U01 AG071448 (Hummer) and U01AG071450 (Aiello and Hummer) at the University of North Carolina at Chapel Hill. Add Health was designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris at the University of North Carolina at Chapel Hill.

In the Section 7: Household Roster example below, the logic for the multiple-choice gender identity question is the same (includes: ask Q6A), and each specific multiple-choice option is not referenced in the logic itself (does not include: ask Q6Aa, ask Q6Ab, ask Q6Ac, etc.).

M1HR06AA: H406AA HH MEMBER 1-MALE- M1	
M1HR06AA	Q6A. What is [...]'s current gender identity? Select all that apply. -- [Member #1] Q6Aa. Male
Logic	if Q4A=1 or 2, or Q3a=1, ask Q6A *Read "ask Q6A" as ask all Q6A multiple-choice questions (Q6Aa-Q6Ag)*

M1HR06AB: H406AB HH MEMBER 1-FEMALE- M1	
M1HR06AB	Q6A. What is [...]'s current gender identity? Select all that apply. -- [Member #1] Q6Ab. Female
Logic	if Q4A=1, 2 or Q3a=1, ask Q6A *Read "ask Q6A" as ask all Q6A multiple-choice questions (Q6Aa-Q6Ag)*

Response Codes and Labels

In the data set, valid responses include a categorical or numerical response (write-ins are not included in the data set). The response labels are included in the Codebooks. The special codes assigned are in Table 9, and include codes for missing, legitimate skips, don't know, and request access (for raw data of sensitive variables).

Missings Include Partial Completes

Seventy-seven respondents who partially completed the survey after Section 11: Mental Health were included in the final data set. Missing "." does not differentiate between partial completes and a respondent purposefully skipping the question. Therefore, the missingness rate may be increasingly inflated as the survey progresses for questions without skip patterns.

The Sexual Orientation/Gender Identity, Socioeconomic Status, and Health Across the Life Course (SOGI-SES) Study is co-directed by Principal Investigators Carolyn T. Halpern at the University of North Carolina at Chapel Hill and Kerith J. Conron at the Williams Institute, UCLA and is funded by the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (NICHD) and the National Institute on Minority Health and Health Disparities under grants R01 HD087365 and R01 HD087365-03S1. SOGI-SES is an Add Health ancillary study that includes a subset of Add Health respondents.

Waves I-V of Add Health were funded by grant P01 HD31921 (Harris) from the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (NICHD), with cooperative funding from 23 other federal agencies and foundations. Add Health is currently directed by Robert A. Hummer and funded by the National Institute on Aging cooperative agreements U01 AG071448 (Hummer) and U01AG071450 (Aiello and Hummer) at the University of North Carolina at Chapel Hill. Add Health was designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris at the University of North Carolina at Chapel Hill.

Table 9: Special Response Codes

Category	Code	Notes
Missing	.*	Could be: - Question received but not answered. - Questions not received because respondent had stopped taking the survey (partial completes).
Legitimate skips	97/997/9997/99997	Question not received
Don't know Codebook Type: Code	3	Applies to categorical responses: 1 = Yes, 2 = No, 3 = Don't know
Don't know Codebook Type: Numeric (Double)	98/998/9998	Applies to write-in or drop-down numeric responses [Codebook Type: Numeric (Double)] such as age
Request access	8888	Raw data of Sensitive variables are not available in the Main data set. Access can be requested.

* Missings include respondents who partially completed the survey (n=77)

Gender Identity and CISFLAG (for internal logic only) Variables

Gender identity variables were determined to be sensitive variables. The raw gender identity variables are only available in the Sensitive data set.

A dummy variable flag, CISFLAG, was added to Section 4: Gender Identity (GI) for internal skip logic purposes. This variable is not intended as a cisgender/transgender definition to use as a demographic or in analyses. Each user may define cisgender and transgender as appropriate for their research questions.

For internal logic purposes, cisgender (CISFLAG=1) is defined as follows: if the gender identity response includes response "E," "I am not sure of my gender identity (I am "questioning" my gender identity)" or response "F," "I do not know what this question is asking" or sex-assigned at birth and gender identity are concordant. Otherwise, CISFLAG=0, indicating the respondent was coded as non-cisgender for internal skip logic only.

Special Sections:

Section 7: Household Roster (HR)

Position of Household Member: Capital letter corresponds to the 1st, 2nd, 3rd, etc. household member reported. For example, see naming convention below for household member age question.

- Household member #1 (Q5A)
- Household member #2 (Q5B)
- Household member #12 (Q5L)

Section 8: Non-Household Child Roster (CR)

Position of Non-Household Child: Capital letter corresponds to the 1st, 2nd, 3rd, etc. non-household child reported. For example, see naming convention below for age question.

The Sexual Orientation/Gender Identity, Socioeconomic Status, and Health Across the Life Course (SOGI-SES) Study is co-directed by Principal Investigators Carolyn T. Halpern at the University of North Carolina at Chapel Hill and Kerith J. Conron at the Williams Institute, UCLA and is funded by the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (NICHD) and the National Institute on Minority Health and Health Disparities under grants R01 HD087365 and R01 HD087365-03S1. SOGI-SES is an Add Health ancillary study that includes a subset of Add Health respondents.

Waves I-V of Add Health were funded by grant P01 HD31921 (Harris) from the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (NICHD), with cooperative funding from 23 other federal agencies and foundations. Add Health is currently directed by Robert A. Hummer and funded by the National Institute on Aging cooperative agreements U01 AG071448 (Hummer) and U01AG071450 (Aiello and Hummer) at the University of North Carolina at Chapel Hill. Add Health was designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris at the University of North Carolina at Chapel Hill.

- Non-household child #1 (Q4A)
- Non-household child #2 (Q4B)
- Non-household child #12 (Q4L)

Section 9: Children's Healthcare Needs

Position of Child with Healthcare Needs: Capital letter corresponds to the 1st, 2nd, 3rd, etc. child with healthcare needs reported. For example, see naming convention below for medicine prescribed by a doctor question.

- Child #1 (Q3A)
- Child #2 (Q3B)
- Child #12 (Q3L)
- Child #24 (Q3X)

To link the individual children from the Rosters to the Children's Healthcare Needs section, Children's Healthcare Needs individuals 1-12 correspond to the 1-12 positions in the Household Roster. Children's Healthcare Needs individuals 13-24 correspond to the 1-12 positions in the Non-Household Child Roster. For example, Child #6 in the Children's Healthcare Needs refers to Member #6 in the Household Roster. Child #16 in the Children's Healthcare Needs refers to the NH Child #4 in the Non-Household Child Roster.

Section 10: Romantic Partner (P)

Type of Romantic Partner: If a respondent did not have a household partner and reported multiple non-household partners, a capital letter was added after the question number where A indicates the reported main partner, B indicates the first current partner listed who is not considered their main partner, and C indicates the second partner listed for both those who completed the Main partner or Current partner #1 questions. For example, see the question naming convention below for how long they have been in a relationship with the designated partner.

- A: Main partner (Q30A)
- B: Current partner #1 (Q30B)
- C: Current partner #2 (Q30C)

Race/ethnicity of Romantic Partner:

- Lowercase letters indicate race/ethnicity questions: a-g
- Numbers indicate race/ethnicity sub-questions based on race/ethnicity responses.

Analyzing the Data

Using SOGI-SES Analysis Weights and Add Health Strata and Cluster Variables

The SOGI-SES eligible sampling frame consisted of living Add Health Wave V respondents with Wave I sampling weights (n=11,712).

The Sexual Orientation/Gender Identity, Socioeconomic Status, and Health Across the Life Course (SOGI-SES) Study is co-directed by Principal Investigators Carolyn T. Halpern at the University of North Carolina at Chapel Hill and Kerith J. Conron at the Williams Institute, UCLA and is funded by the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (NICHD) and the National Institute on Minority Health and Health Disparities under grants R01 HD087365 and R01 HD087365-03S1. SOGI-SES is an Add Health ancillary study that includes a subset of Add Health respondents.

Waves I-V of Add Health were funded by grant P01 HD31921 (Harris) from the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (NICHD), with cooperative funding from 23 other federal agencies and foundations. Add Health is currently directed by Robert A. Hummer and funded by the National Institute on Aging cooperative agreements U01 AG071448 (Hummer) and U01AG071450 (Aiello and Hummer) at the University of North Carolina at Chapel Hill. Add Health was designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris at the University of North Carolina at Chapel Hill.

Due to the complicated sampling designs of SOGI-SES and Add Health, observations are not independent, complicating statistical analyses. To analyze the data correctly, special survey statistical packages and the correct sampling weights must be identified and utilized. See Guidelines for Analyzing Add Health Data (doi.org/10.17615/C6BW8W) for more information about elements needed by the survey software packages as well as common errors to avoid when analyzing the Add Health data.

A summary of the attributes of the Add Health Sampling Design paired with the SOGI-SES cross-sectional weight that may be used for analyses is included in Table 10 with the variable name and location of access. Unique to the SOGI-SES study is a SOGI-SES cross-sectional analysis weight, which is used when the outcome variable is in the SOGI-SES data set. The variable name is M1WEIGHT and is available in the Main data set and Codebook under Section A: Cross-Sectional Analysis Weight and Survey Information (WS).

Table 10: Attributes of Add Health Sampling Design

Design Attribute	Variables in Add Health and SOGI-SES Data Used to Adjust for the Sampling Design	Study	Data Access Location	Variable Name
Stratification	POSTRATIFICATION VARIABLE: Census Region	Add Health	Wave I Weights Codebook	REGION
Clustering of Students	PRIMARY SAMPLING UNIT VARIABLE: School Identification Variable	Add Health	Wave I Weights Codebook	PSUSCID
Unequal Probability of Selection	ANALYSIS WEIGHTS: SOGI-SES cross-sectional analysis weights	SOGI-SES	SOGI-SES Codebook	M1WEIGHT

Examples of basic set up in Stata and SAS for outcomes in the SOGI-SES data set:

- Stata: svyset psuscid [pw = m1weight], strata(region)
- SAS:

```
proc surveymeans;
var sogi-var;
cluster PSUSCID;
strata REGION;
weight M1WEGHT;
run;
```

Merging data sets

To use the Add Health data variables with SOGI-SES data, it is critical to merge the data sets correctly, using the respondent's AID, the Add Health respondent's assigned ID that is applied across all Add Health waves and SOGI-SES. Use of pregnancy, childbirth, and partner data from Add Health surveys requires use of additional IDs. Please see the Add Health website for information about merging files.

The Sexual Orientation/Gender Identity, Socioeconomic Status, and Health Across the Life Course (SOGI-SES) Study is co-directed by Principal Investigators Carolyn T. Halpern at the University of North Carolina at Chapel Hill and Kerith J. Conron at the Williams Institute, UCLA and is funded by the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (NICHD) and the National Institute on Minority Health and Health Disparities under grants R01 HD087365 and R01 HD087365-03S1. SOGI-SES is an Add Health ancillary study that includes a subset of Add Health respondents.

Waves I-V of Add Health were funded by grant P01 HD31921 (Harris) from the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (NICHD), with cooperative funding from 23 other federal agencies and foundations. Add Health is currently directed by Robert A. Hummer and funded by the National Institute on Aging cooperative agreements U01 AG071448 (Hummer) and U01AG071450 (Aiello and Hummer) at the University of North Carolina at Chapel Hill. Add Health was designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris at the University of North Carolina at Chapel Hill.

Cross-sectional analyses can be conducted with the SOGI-SES cross-sectional analysis weights (M1WEIGHT). Because SOGI-SES longitudinal weights were not created, depending on your outcome variable and other components, longitudinal analyses with repeated outcome measures across Add Health survey waves and the SOGI-SES survey may not be possible. See Guidelines for Analyzing Add Health Data (doi.org/10.17615/C6BW8W) for guidance and examples of when to use cross-sectional and longitudinal weights. While the Add Health Guidelines that were developed in 2020 did not specifically discuss SOGI-SES, the general principles apply. Below are examples of explanations found in the Guidelines for Analyzing Add Health Data (doi.org/10.17615/C6BW8W).

Resources

- Add Health - https://addhealth.cpc.unc.edu/
  - Documentation tab: SOGI-SES Codebooks, Index, User Guide, and Add Health materials.
  - Data tab: information about how to access all Add Health data sets including SOGI-SES
- CPC Data Portal – https://data.cpc.unc.edu

Acknowledgements

The SOGI-SES and Add Health contracts and data use agreements require that the following be included in each written report or other publication based on analysis of the SOGI-SES and Add Health data.

The Sexual Orientation/Gender Identity, Socioeconomic Status, and Health Across the Life Course (SOGI-SES) Study is co-directed by Principal Investigators Carolyn T. Halpern at the University of North Carolina at Chapel Hill and Kerith J. Conron at the Williams Institute, UCLA and is funded by the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD) and the National Institute on Minority Health and Health Disparities under grants R01 HD087365 and R01 HD087365-03S1. SOGI-SES is an Add Health ancillary study that includes a subset of Add Health respondents.

Add Health was funded by grant P01 HD31921 (Harris) from the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD), with cooperative funding from 23 other federal agencies and foundations. Add Health is currently directed by Robert A. Hummer and funded by the National Institute on Aging cooperative agreements U01 AG071448 (Hummer) and U01AG071450 (Aiello and Hummer) at the University of North Carolina at Chapel Hill. Add Health was designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris at the University of North Carolina at Chapel Hill.

Research reported in this publication was supported by the Carolina Population Center funded by NICHD of the National Institutes of Health under award number P2C HD050924.

Special Thank You to the Following People for their Contributions:

Thank you to the work and expertise contributed by the staff of RTI International including Gretchen McHenry for her work on the cognitive interviews, Gretchen McHenry and Lilia Filippenko for helping design and program the online survey, Paul Biemer and Dan Liao for developing the cross-sectional weights and conducting the non-bias response analysis, Chris Carson for recruitment, data, and data transfer activities. Thank you to Brian Burke who managed the project.

The Sexual Orientation/Gender Identity, Socioeconomic Status, and Health Across the Life Course (SOGI-SES) Study is co-directed by Principal Investigators Carolyn T. Halpern at the University of North Carolina at Chapel Hill and Kerith J. Conron at the Williams Institute, UCLA and is funded by the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (NICHD) and the National Institute on Minority Health and Health Disparities under grants R01 HD087365 and R01 HD087365-03S1. SOGI-SES is an Add Health ancillary study that includes a subset of Add Health respondents.

Waves I-V of Add Health were funded by grant P01 HD31921 (Harris) from the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (NICHD), with cooperative funding from 23 other federal agencies and foundations. Add Health is currently directed by Robert A. Hummer and funded by the National Institute on Aging cooperative agreements U01 AG071448 (Hummer) and U01AG071450 (Aiello and Hummer) at the University of North Carolina at Chapel Hill. Add Health was designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris at the University of North Carolina at Chapel Hill.

Thank you to Lixin Qu, the SOGI-SES programmer, data manager, coder and analyst.

Thank you to the Add Health team for their ongoing support throughout the study process and for managing the study materials and data after data release. Special thanks to Julie Kouzoukas, Sarah Dean, Sydney Will, and Mary Jane Hill.

Thank you to our Scientific Advisors, Drs. Lee Badgett (University of Massachusetts Amherst), Randy Sell (Drexel University), and Ilan Meyer (University of California, Los Angeles), for their feedback and assistance with survey development.

Thank you to Lynn Balen and Pooja Deshpande for their contributions and review of the Codebook.

The Sexual Orientation/Gender Identity, Socioeconomic Status, and Health Across the Life Course (SOGI-SES) Study is co-directed by Principal Investigators Carolyn T. Halpern at the University of North Carolina at Chapel Hill and Kerith J. Conron at the Williams Institute, UCLA and is funded by the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (NICHD) and the National Institute on Minority Health and Health Disparities under grants R01 HD087365 and R01 HD087365-03S1. SOGI-SES is an Add Health ancillary study that includes a subset of Add Health respondents.

Waves I-V of Add Health were funded by grant P01 HD31921 (Harris) from the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (NICHD), with cooperative funding from 23 other federal agencies and foundations. Add Health is currently directed by Robert A. Hummer and funded by the National Institute on Aging cooperative agreements U01 AG071448 (Hummer) and U01AG071450 (Aiello and Hummer) at the University of North Carolina at Chapel Hill. Add Health was designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris at the University of North Carolina at Chapel Hill.