

WELCOME
TO THE 2024
NDACAN
SUMMER
TRAINING
SERIES!

- The session will begin at 12pm Eastern.
- Please submit questions to the Q&A box.
- This session is being recorded.

NDACAN SUMMER TRAINING SERIES: BEST PRACTICES IN THE USE OF NDACAN DATA

National Data Archive on Child Abuse and Neglect

Cornell University & Duke University

APPROACHING NSCAW III
FOR EXPERIENCED AND
NEW USERS

AUGUST 7, 2024



Children's Bureau

An Office of the Administration for Children & Families

NDACAN SUMMER TRAINING SERIES SCHEDULE

- July 10 — NCANDS: Strengths & Limitations
- July 17 — Assessing Reporting Issues in NCANDS & AFCARS
- July 24 — AFCARS: Strengths & Limitations
- July 31 — Survey Design & Using Weights
- August 7 — NSCAW III for Experienced & New Users
- August 14 — NYTD: Strengths & Limitations

SESSION AGENDA

- National Survey of Child and Adolescent Well-Being (NSCAW) Background
- Analysis File Creation
- Variable Creation
- Quality Control of Created Variables
- Data Management Best Practices
- Analysis Reminders

NSCAW BACKGROUND

NSCAW COHORT STUDY DESIGNS

Study Characteristic	NSCAW I	NSCAW II	NSCAW III
Completed Interviews	5,501 LTFC: 727*	5,872	3,298
Baseline Age of Children	0 – 14	0 – 17.5	0 – 17.5
Baseline Interview Field Period	11/1999 to 12/2000	3/2008 to 4/2009	8/2017 to 3/2022**

* Long-term foster care (LTFC) survey only conducted in NSCAW I

** Because of the COVID-19 pandemic, baseline data collection was paused from March 2020 until May 2021

INSTRUMENTATION SUMMARY

Instrument/ Respondent	NSCAW I				NSCAW II			NSCAW III	
	Baseline	Wave 2	Wave 3,4	Wave 5	Baseline	Wave 2	Wave 3	Baseline	Wave 2
Child/Youth	X		X	X	X	X	X	X	X
Young Adult (18+ years and emancipated)				X		X	X		X
Caregiver	X	X	X	X	X	X	X	X	X
Teacher	X		X	X	X	X			
Investigative Caseworker	X				X			X	
Services Caseworker		X	X	X		X	X		X
Local Agency Director	X				X				

Notes:

1. No Teacher survey in NSCAW III
2. In NSCAW III, Agency Director survey completed as part of separate set of workforce surveys

NSCAW DATA AND DOCUMENTATION

- The NSCAW I, NSCAW II, and NSCAW III baseline dataset and documentation is available to qualified researchers. Please contact NDACAN if you are interested in receiving these data and documentation.
- Links to additional NSCAW documentation:
 - **Crosswalk of NSCAW measures:**
Crosswalk of Constructs and Measures Across Three Cohorts of the National Survey of Child and Adolescent Well-Being | The Administration for Children and Families (hhs.gov)
<https://www.acf.hhs.gov/opre/report/crosswalk-constructs-and-measures-across-three-cohorts-national-survey-child-and>
 - **NSCAW III Baseline Introductory Report:**
NSCAW III Baseline Report (2017-2022): Introduction to NSCAW III | The Administration for Children and Families (hhs.gov)
<https://www.acf.hhs.gov/opre/report/nscaw-iii-baseline-report-2017-2022-introduction-nscaw-iii>
 - **OPRE NSCAW webpage, including links to analysis products from all three NSCAW cohorts:**
National Survey of Child and Adolescent Well-Being (NSCAW) | The Administration for Children and Families (hhs.gov)
<https://www.acf.hhs.gov/opre/project/national-survey-child-and-adolescent-well-being-nscaw-1997-2014-and-2015-2024>

ANALYSIS FILE CREATION

NSCAW III DATA ARE COMPLEX

- Complex sample design (e.g., weights, clusters, subpopulations)
- Multiple surveys with complex skip patterns
 - Child, caregiver and caseworker instruments
 - Record may not have all modules, instruments, or items
- Many derived variables

Detailed information is provided in the Data File User's Manual (DFUM), appendices, and electronic codebook.

ANALYSIS FILE CONSIDERATIONS

- Which instrument/module(s) contain the data of interest?
- What population are you focusing on?
- What derived variables are available?
- What variables do you need to create?

WHAT INSTRUMENT/MODULE(S) CONTAIN
THE DATA OF INTEREST?

Examples of Content area	Child	Caseworker	Caregiver
Substance use	SA		
Maltreatment	CM	AA	DS
Special education status and services		SC	HS
Emotional/behavioral health problems	YB		TC, BC
Services received	SV	SP, SC	SR

WHAT POPULATION ARE YOU FOCUSING ON?

- Some constructs are only asked of children of certain ages
- Assessments are done on specific ages
- Do not subset the analysis file to your population of interest
- Do keep the minimum data necessary

WHAT DERIVED VARIABLES ARE AVAILABLE?

- Appendix II of the Data File User's Manual (DFUM)
- Scales and assessments for Child and Caregiver instruments
- Variables created from more than one question within an instrument
 - e.g., caregiver highest degree
- Variables created from multiple instruments
 - e.g., child age

WHAT VARIABLES DO YOU NEED TO CREATE?

- Is the information available in more than one place?
- Any skip patterns leading into the question of interest?
- Collapse categories?

VARIABLE CREATION

WHO SAID WHAT?

- Multiple informants are asked similar questions are asked about the child/youth
 - Y: youth, P: caregiver, C: caseworker.
- Which do you prioritize?

WHAT DOES IT MEAN IF A VALUE IS MISSING?

- Many reasons why a response to a variable is not present
 - Skipped due to a skip pattern
 - Respondent didn't know or refused to answer to the question
 - Respondent had only a partial interview

MISSING VALUES IN THE NSCAW III DATA

Missing codes for each variable are listed in the NSCAWIII documentation

PSR14 Have you ever received any job-related services?

Text of this Question or Item

Have you ever received any job-related services?

1 = YES

2 = NO

Percent	N	Value	Label
5.99	15810.12185	-7	Legitimate skip
3.62	95505.15368	-6	Non-interview
0.49	12996.675944	-1	Dont know
5.87	154758.0562	1	YES
84.02	2216308.9923	2	NO
100.0	2637679		Total

Properties

Data type: numeric

Format: tyesnoc

VARIABLE CREATION: EXAMPLE

We will use
dummy data!

What percentage
of children were
referred for dental
care?

NSCAW III DOCUMENTATION

CCI17ATG_28 Services provided-Option 28 - Dental Exam

Text of this Question or Item

USE CARD 7. What kind of services? CODE ALL THAT APPLY

Value	Label
N	NO
S	Legitimate skip
U	Non-interview
Y	YES

INITIAL FREQUENCY

Services provided-Option 28 - Dental Exam .

cci17atg_28	Weighted Frequency	Percent
N	532699	20.1958
S	1170960	44.3936
U	587789	22.2843
Y	346230	13.1263
Total	2637679	100.0000

FREQUENCY OF THOSE WHO ANSWERED

```
if CCI17ATG_28='Y' then Dental=1; /*dental*/  
else if CCI17ATG_28='N' then Dental=0; /*no dental*/
```

Table of Dental		
Dental	Weighted Frequency	Percent
No	532699	60.6077
Yes	346230	39.3923
Total	878930	100.0000
Frequency Missing = 2047		

SECOND LOOK AT MISSING CODES

Value	Label
N	NO
S	Legitimate skip
U	Non-interview
Y	YES

Services provided-Option 28 - Dental Exam .

cci17atg_28	Weighted Frequency	Percent
N	532699	20.1958
S	1170960	44.3936
U	587789	22.2843
Y	346230	13.1263
Total	2637679	100.0000

QUESTION THAT CAUSED THE SKIP

CCI16A Services provided/arranged for family

Text of this Question or Item

Regardless of the case decision of the investigation/assessment, have any services been referred for, provided to, or arranged for the family? Referring the family for services includes suggesting to the client that services may be needed, or giving the client provider contact information. Arranging services for the family includes contacting a provider, completing the necessary paperwork, and/or making an appointment.

FINAL FREQUENCY

```
if CCI17ATG_28='Y' then Dental=1; /*dental*/  
else if CCI17ATG_28='N' OR CCI16A=2 then Dental=0; /*no dental*/
```

Table of Dental		
Dental	Weighted Frequency	Percent
No	1687117	82.9724
Yes	346230	17.0276
Total	2033348	100.0000

Frequency Missing = 714

QC OF CREATED VARIABLES

QC

- All created variables should be QC'ed
 - Code review
 - Individual record review for sample of records
 - Crosstabs of raw and created variables

QC CROSSTAB CODE

```
proc freq data=&dsout;  
tables cci16a*cci17atg_28*dental/list missing ;  
run;
```

QC CROSSTAB OUTPUT

cci16a	cci17atg_28	Dental	
Legitimate skip	S	.	
Non-interview	U	.	
Don't know	N	No	
YES	N	No	
YES	Y	Yes	
NO	S	No	

DATA MANAGEMENT BEST PRACTICES

BEST PRACTICES

- Document all variable creation in a document separate from program
- Create a permanent dataset specific to your analysis task (i.e., paper, report, etc.)
- Minimum data necessary
- QC creation of variables
 - Pay close attention to missing values
- Consider a code repository (ie, macro program) for variables used in multiple tasks

ANALYSIS REMINDERS

NSCAW III ANALYSIS REMINDERS

- Weighting
- Subpopulation analysis
- Unreliable estimates (using RSE)
- Suppression due to small cell size in tables

WEIGHTING

- Use software package to adjust for sampling
 - e.g., SUDAAN, SAS Survey Sampling procedures, etc.
- Use weight for all means, proportions, percentages, or other types of estimates
 - NANALWT
- Use PSU identifier to generate adjusted variance and standard errors of estimates
 - NSCAWPSU

```
proc surveyfreq;  
weight nanalwt;  
cluster nscawpsu;  
tables dental/cl;  
run;
```

WEIGHTING EXAMPLE: SURVEY ESTIMATES

Table of Dental							
Dental	Frequency	Weighted Frequency	Std Err of Wgt Freq	Percent	Std Err of Percent	95% Confidence Limits for Percent	
No	2105	1687117	197241	82.9724	1.2920	80.3881	85.5567
Yes	479	346230	38464	17.0276	1.2920	14.4433	19.6119
Total	2584	2033348	225149	100.0000			

Frequency Missing = 714

SUBPOPULATION ANALYSIS

- Do not subset analysis file
- Can result in incorrect variance and standard errors
- Use SUBPOPN in SUDAAN , DOMAIN statement in SAS SURVEYMEANS or equivalent statement in Stata/SPSS/R

SUBPOPULATION EXAMPLE: FEMALES AGED 17

INCORRECT

- Subset file to records of interest

```
IF CHDAGEY = 17 AND  
CHDBIRTHSEX = 2;
```

- Use WHERE statement in procedure

```
WHERE CHDAGEY = 17 AND  
CHDBIRTHSEX = 2;
```

CORRECT

- Keep all records but create a variable (if necessary) to identify subpopulation

```
FEMALE_AGE17 = (CHDAGEY = 17  
AND CHDBIRTHSEX = 2);
```

QC FEMALE_AGE17 CREATION

chdAgeY	CHDBIRTHSEX	FEMALE_AGE17
0	Male	0
0	Female	0
1	Male	0
1	Female	0
2	Male	0
2	Female	0
3	Male	0

• • •

16	Male	0
16	Female	0
17	Male	0
17	Female	1

SUBPOPULATION EXAMPLE: FEMALES AGED 17 (CTD.)

INCORRECT

```
PROC SURVEYFREQ;  
WHERE FEMALE_AGE17=1;  
CLUSTER NSCAWPSU;  
WEIGHT NANALWT;  
TABLES EVERSEX / CL;  
RUN;
```

CORRECT

```
PROC SURVEYFREQ;  
CLUSTER NSCAWPSU;  
WEIGHT NANALWT;  
TABLES FEMALE_AGE17 * EVERSEX / ROW  
CL;  
RUN;
```

SUBPOPULATION EXAMPLE: INCORRECT OUTPUT

The SURVEYFREQ Procedure

Data Summary

Number of Clusters	35
Number of Observations	58
Sum of Weights	31493.5324



Child has ever had sex

eversex	Frequency	Weighted Frequency	Std Err of Wgt Freq	Percent	Std Err of Percent	95% Confidence Limits for Percent	
No	15	7018	2607	22.61	7.54	7.29	37.93
Yes	41	24023	5908	77.39	7.54	62.07	92.71
Total	56	31041	6622	100.00			

Frequency Missing = 2

SUBPOPULATION EXAMPLE: CORRECT OUTPUT

The SURVEYFREQ Procedure

Data Summary

Number of Clusters	61
Number of Observations	3298
Sum of Weights	2637679

Table of FEMALE_AGE17 by eversex

FEMALE_AGE17	eversex	Frequency	Weighted Frequency	Std Err of Wgt Freq	Row Percent	Std Err of Row Percent	95% Confidence Limits for Row Percent	
0	No	611	536177	62216	50.27	2.56	45.14	55.40
	Yes	678	530473	62804	49.73	2.56	44.60	54.86
	Total	1289	1066650	112423	100.00			
1	No	15	7018	2706	22.61	7.49	7.62	37.60
	Yes	41	24023	6451	77.39	7.49	62.40	92.38
	Total	56	31041	7432	100.00			
Total	No	626	543195	62863				
	Yes	719	554496	65335				
	Total	1345	1097691	115548				

Frequency Missing = 1953

SUBPOPULATION EXAMPLE: COMPARING ESTIMATES

INCORRECT

eversex		Frequency	Weighted Frequency	Std Err of Wgt Freq	Percent	Std Err of Percent	95% Confidence Limits for Percent	
Yes		41	24023	5908	77.39	7.54	62.07	92.71

CORRECT

FEMALE_AGE17	eversex	Frequency	Weighted Frequency	Std Err of Wgt Freq	Row Percent	Std Err of Row Percent	95% Confidence Limits for Row Percent	
1	Yes	41	24023	6451	77.39	7.49	62.40	92.38

UNRELIABLE ESTIMATES

- Relative Standard Error
 - $(\text{SE of point estimate}) / (\text{point estimate})$, expressed as a percentage

Eversex	Percent	SE of Percent	RSE calculation	RSE as a %
No	22.61	7.54	$7.54 / 22.61$	33%
Yes	77.39	7.54	$7.54 / 77.39$	10%

 **Unreliable**

- Any estimate with RSE greater or equal to 25% should be noted
 - Footnote, symbol, etc.

SUPPRESSION DUE TO SMALL CELL SIZES

- Any estimate based on an **unweighted** N less than 11
 - Numerator or denominator for percentage estimates
- Suppressed from text and tables

	Peer Acceptance Total (N= 1,148)	
	<i>N</i>	Mean
Race		
American Indian/Alaskan Native	52	3.3
Asian	--	--
Black	297	3.8
Native Hawaiian/Pacific Islander	--	--
White	634	3.2
Two or more races	150	3.5

Cells are suppressed (--) based on cell count less than 11.

Note: Not actual NSCAW estimates. Dummy data was used.

QUESTIONS?

MARIANNE KLUCKMAN
RESEARCH SYSTEMS
PROGRAMMER/ANALYST

MKLUCKMAN@RTI.ORG

NEXT WEEK...

**August 14th, 2024 at
12pm (Eastern)**

Presenters:

Sarah Sernaker and Tammy White

Topic:

NYTD Strengths and Limitations