



# Introduction to the Research and Development Survey and Calibration Approaches for Panel Surveys

**Katherine Irimata**

November 3, 2021

# National Center for Health Statistics

- The National Center for Health Statistics (NCHS) is the principal health statistics agency in the United States
- NCHS monitors the nation's health by collecting, analyzing, and disseminating health data to:
  - Document the health status of the U.S. population and selected subgroups
  - Identify disparities in health status and use of health care
  - Inform and evaluate health policies
  - Monitor trends in health indicators

# NCHS Data Systems

- NCHS collects health data using various data systems:



National Vital  
Statistics  
System



National  
Health  
Interview  
Survey



National  
Health and  
Nutrition  
Examination  
Survey



National  
Health Care  
Surveys



National  
Survey of  
Family  
Growth



# New Data Sources: Panel Surveys

- Panels contain a group of potential survey respondents maintained by private companies (e.g., Gallup, NORC, IPSOS)
- Panelists can be recruited passively using a non-probability sample (opt-in) or actively based on statistical sampling (recruited)
- Many panel surveys utilize web and phone modes
- Relatively faster data collection and lower cost compared to in-person household surveys
- Limitations include lower sample sizes, higher non-response bias, and coverage bias

# Research and Development Survey

- NCHS started the Research and Development Survey (RANDS), an ongoing, periodic set of probability-sampled commercial panel surveys, in 2015
- Established in response to the growing interest in alternative survey modes and the maturation of relatively inexpensive, commercial survey panels
- Designed for methodological purposes
  - Measurement error: supplement survey and questionnaire evaluation efforts
  - Estimation: explore ways to integrate data from commercial survey panels with high-quality data collections, including calibration

# Outline

- RANDS
  - Questionnaire
  - Survey Design
  - Timeline
  - Completed Rounds
- Calibration
  - Background
  - Calibration in RANDS data
- Summary

**RANDS**

# Questionnaire

- Questions focus on a range of health-related topics including chronic conditions, access to healthcare and utilization, and opioid use
- Rounds may have specific topic focus (e.g., RANDS 4 – disability, etc.)
- Most questions in completed rounds have been from the National Health Interview Survey (NHIS) questionnaire
- Embedded probe questions are also used to examine respondent question interpretation



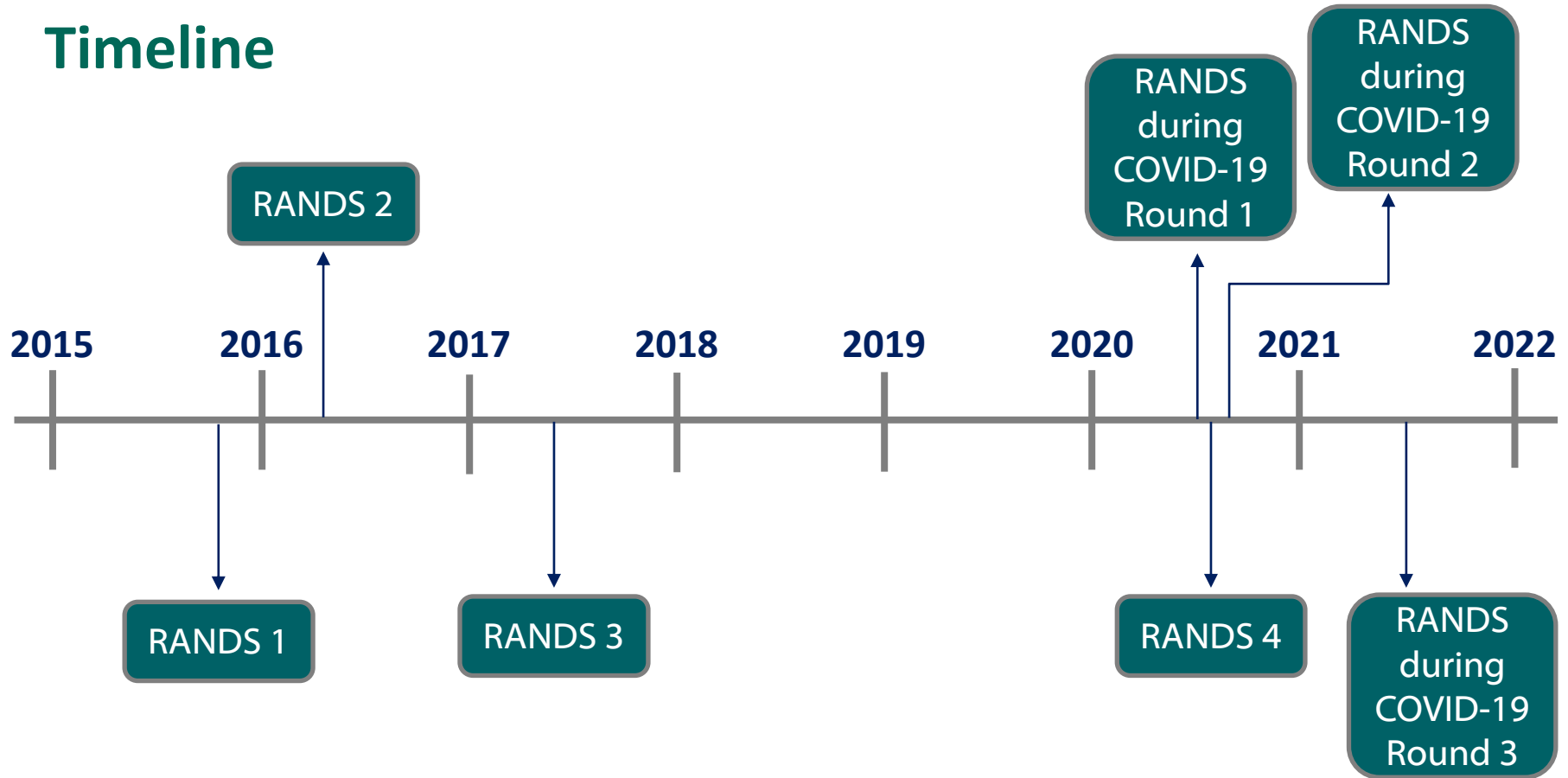
# Survey Design

- Target population includes general adult population of the United States (18 years and older)
- Sample drawn from probability-based sample using stratified sample design
- Sampling strata formed using combinations of demographic variables (such as age group, race/ethnicity, sex, etc.)
- Within strata, sampling performed using simple random sampling
- Panel weights account for selection into the commercial panel and for selection into RANDS, with adjustments for non-response and under coverage of groups on the sampling frame

# Rounds

- RANDS started in 2015 and are numbered according to the order of collection
- Four rounds of regular collection have been conducted (RANDS 1, RANDS 2, RANDS 3, and RANDS 4)
- In response to the COVID-19 pandemic, a special series of RANDS was used to report on the impact of the pandemic termed RANDS during COVID-19
- Three rounds of RANDS during COVID-19 have been conducted (RANDS during COVID-19 Rounds 1, 2, and 3)

# Timeline



# Overview of completed rounds of RANDS

Survey	RANDS 1	RANDS 2	RANDS 3	RANDS 4
Dates	11/2/15-12/9/15	3/29/16-4/13/16	4/11/19-4/24/19	7/17/20-8/24/20
Panel	Gallup	Gallup	AmeriSpeak	AmeriSpeak
Completed Surveys	2,304	2,480	2,646	3,442
Completion Rate	23.5%	30.1%	62.2%	70.0%
Cumulative Response Rate	Unknown	Unknown	18.1%	14.0%
Mode	Web	Web	Web	Web and phone
Survey Focus	Health Conditions and Behaviors	Health Conditions and Behaviors	Disability and Opioids	Disability and Opioids

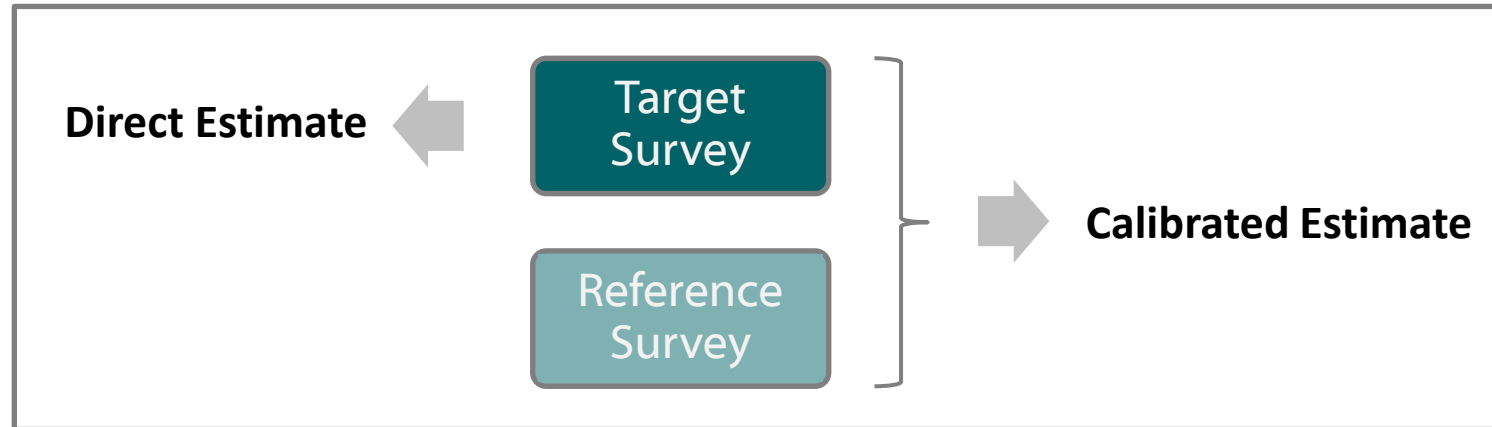
# Overview of completed rounds of RANDS during COVID-19

Survey	RANDS during COVID-19 Round 1	RANDS during COVID-19 Round 2	RANDS during COVID-19 Round 3
Dates	6/9/20-7/6/20	8/3/20-8/20/20	5/17/21-6/30/21
Panel	AmeriSpeak	AmeriSpeak	AmeriSpeak
Completed Surveys	6,800	5,981	5,452
Completion Rate	78.5%	69.1%	69.5%
Cumulative Response Rate	23.0%	20.3%	11.8%
Mode	Web and phone	Web and phone	Web and phone
Survey Focus	COVID-19	COVID-19	COVID-19

**Calibration**

# Background

- Combining web-based panel surveys with high quality reference data can improve estimates of national health outcomes
- Calibration to the reference survey reduces selection bias (Taylor, 2000; Terhanian and Bremer, 2000)



# Calibration Approaches

- Raking
  - Iterative method of adjusting the sample weights to reflect external population counts
  - Specified measure of convergence (e.g., number of iterations, percentage tolerance)
- Propensity score (PS)-based methods (matching, weighting, stratification)
  - Adjusts for nonresponse bias by modeling the probability of inclusion ( $p$ ) in terms of survey covariates
  - Flexibility in model formation (e.g., interactions, higher order terms)



# Considerations

- Variable selection
  - Demographic variables
  - Variables that may identify differences between respondents in the target survey and reference survey
  - Variables associated with the outcome
- Missing data
- Alignment of calibration variables in target and reference surveys

# Data Quality Checks

- Descriptive statistics of weights
  - Compare distributions of weights
  - Identify outliers
- Balance of outcome among covariates in target and reference datasets
- Convergence of calibration method/model
- Assess calibrated estimates and standard errors

# Calibration in RANDS Data

- RANDS panel weights have been calibrated to the NHIS sample adult weights
- Annual NHIS data from corresponding RANDS year used for calibration (or most recently available public data)
- Calibration variables have included selected demographic and health variables
- PS-based methods and raking have been used for calibration research, although raking was used to produce calibrated weights for RANDS during COVID-19
- After calibration, adjusted RANDS estimates are generally closer to NHIS estimates

# Summary

# Summary

- RANDS is ...
  - A series of probability-sampled commercial panel surveys
  - Designed for methodological research purposes, including calibration research to improve estimates of health outcomes
  - Flexible and was quickly adapted to report on the COVID-19 pandemic
- The upcoming session presentations provide specific uses and comparisons from RANDS rounds

<https://www.cdc.gov/nchs/rands>

Katherine Irimata  
kirimata@cdc.gov

For more information, contact CDC  
1-800-CDC-INFO (232-4636)  
TTY: 1-888-232-6348 [www.cdc.gov](http://www.cdc.gov)

The findings and conclusions in this presentation are those of the authors and do not necessarily represent the official position of the National Center for Health Statistics, Centers for Disease Control and Prevention.



# References

- Taylor, H. (2000). Does Internet Research Work? Comparing Online Survey Result with Telephone Survey. *International Journal of Market Research*, 42, 58-63.
- Terhanian, G. and Bremer, J. (2000). Confronting the Selection-Bias and Learning Effects Problems Associated with Internet Research. Research paper: Harris Interactive.