Challenges in Biomonitoring Surveillance Studies: Response Rates for Population-Based Surveys

Brian M. Wells, PhD
Data Quality and Survey Methodology Manager
California Health Interview Survey (CHIS)
UCLA Center for Health Policy Research

Biomonitoring California Scientific Guidance Panel – November 12, 2020

Disclaimer

- The interpretations, conclusions, and views expressed in this presentation are those of the author and do not necessarily represent the UCLA Center for Health Policy Research, the Regents of the University of California, or collaborating organizations or funders.

Biomonitoring, biomeasures, and biomarkers

- **Biomonitoring** assesses human exposure to an environmental chemical by measuring the chemical through blood, urine, saliva, etc.
- **Biomeasures** collectively refers to anthropomorphic measures, physical performance measures, and biological material
  (Jazczak et al., 2009; Sakshaug et al., 2015)
- **Biomarkers** are biological indicators of a process, event, or condition
  - Generally an assay generated from a biomeasure sample

- We **biomonitor** by observing **biomarkers** from **biomeasures**

Biomeasure collection

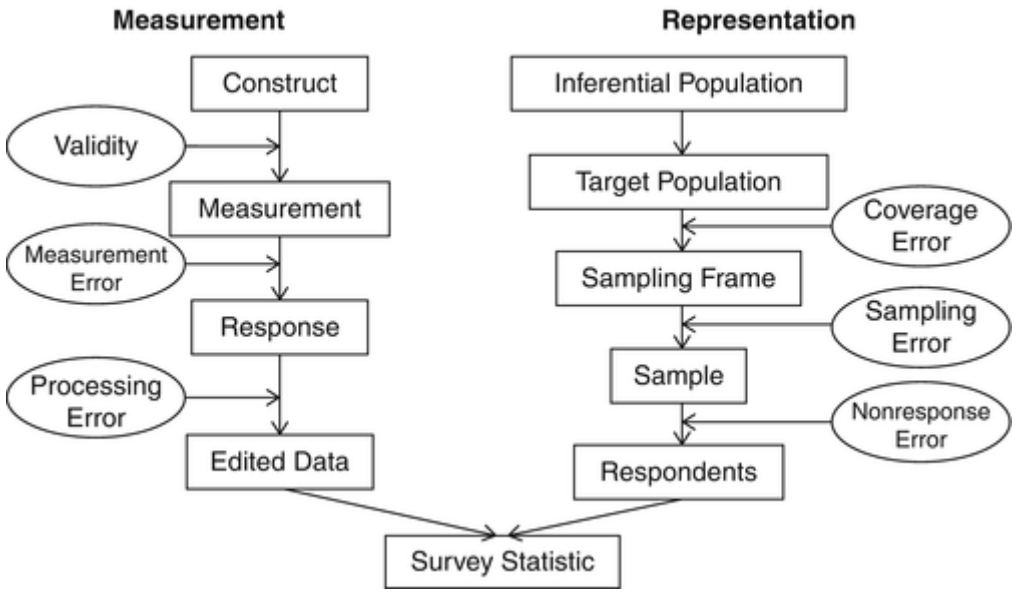
- Biomarkers of interest determine biomeasures of collection
- Biomeasure collection approaches in survey research
  - **Medically-trained nurse or phlebotomist**
    - In-home, at a clinic, or other location
    - Example: National Health and Nutrition Examination Survey (NHANES)
  - **Non-medically-trained interviewer**
    - Limited to minimally invasive collections (Lindau & McDade, 2008)
      - Example: Dried blood spot assay, saliva catch
    - Example: National Social Life, Health, and Aging Project (NSHAP) (Jaszczak et al., 2009)
  - **Self-administered, mail back**
    - Limited to minimally invasive collections
    - Example: Health and Retirement Study (HRS) 2003 Diabetes Study (Weir, 2008)

Biomonitoring and population-based research

- The goal of population-based research is to produce findings generalizable to the target population and can be used for population health surveillance
- Surveys are often used to obtain information from a sample (or random subset) of the population
- A common concern about population-based surveys is representation

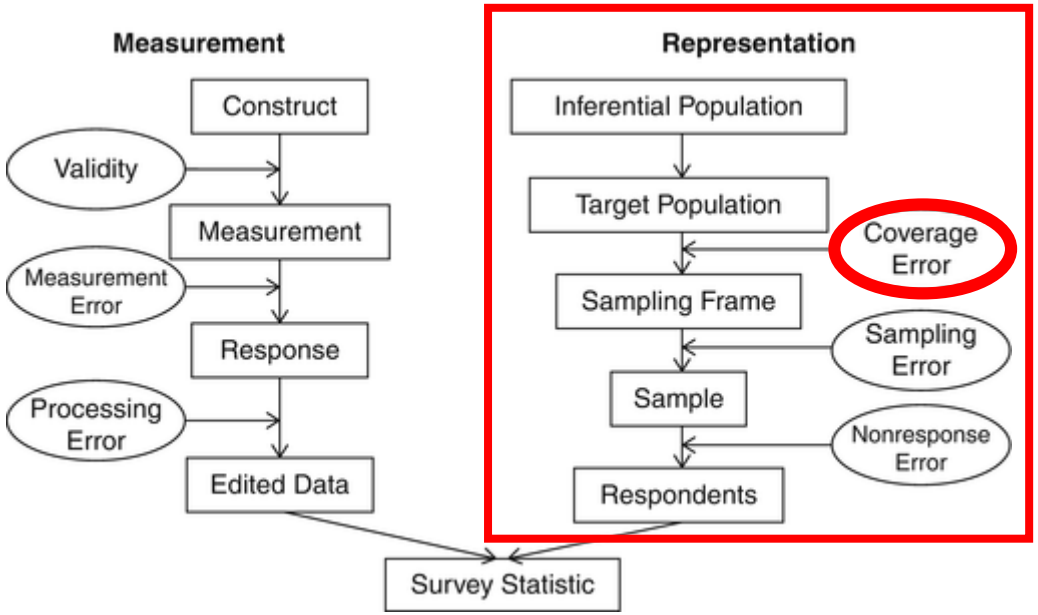
Representation in population-based surveys

- Total Survey Error (TSE) paradigm
  (e.g., Biemer & Lyberg, 2003)
- Multiple sources of error broken into **Measurement** and **Representation**
  - Measurement = Constructs, questions and responses
  - Representation = Respondents and the population of interest
- All surveys have error; our job is to minimize the errors



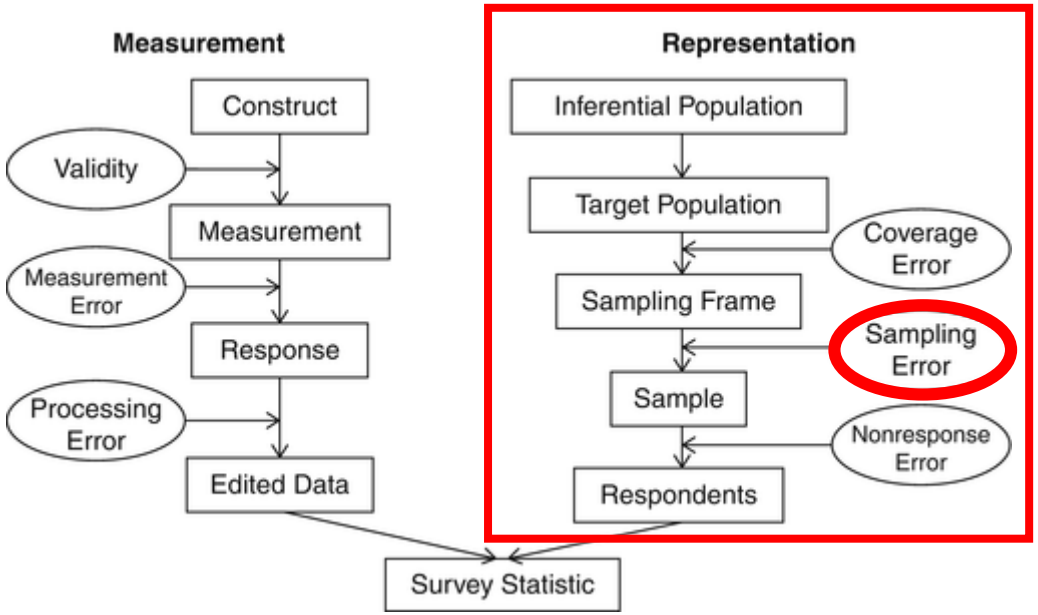
Representation: Coverage error

- Sampling frame is incomplete in regards to the target population
- Most common coverage problem is **under coverage**
- Ex. Random-digit dialing (RDD) → not everyone has a telephone
- Ex. Address-based sampling (ABS) → not everyone has an address



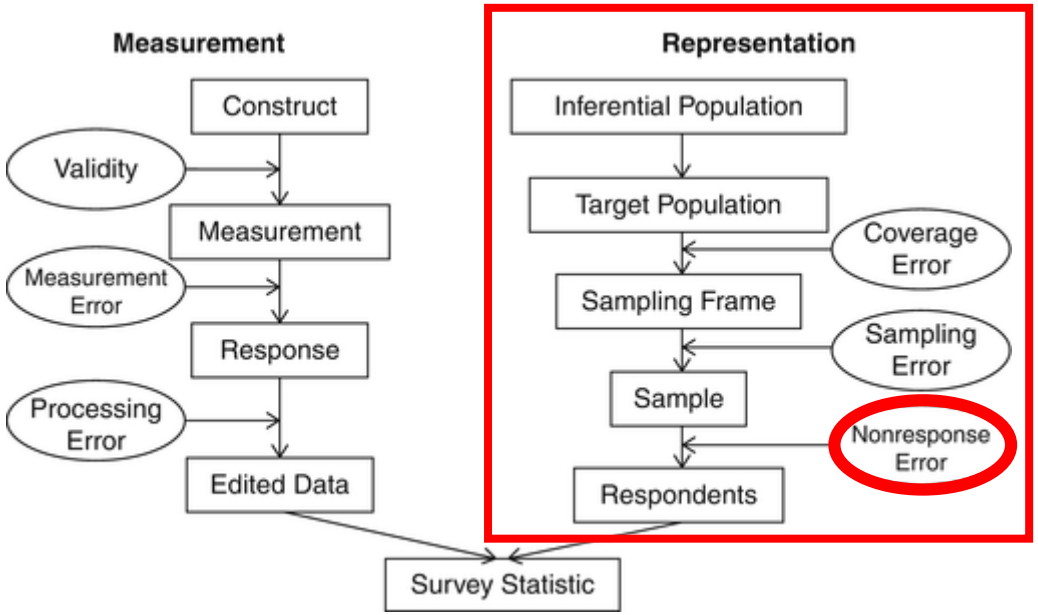
Representation: Sampling error

- How we determine the sample to survey from the sampling frame
- Simple random sample vs. cluster vs. stratified vs. complex vs. something else
- Allows us to quantify variability
  - Primarily affects variances (larger or smaller)
- Non-probability designs is a concern



Representation: Nonresponse error

- Sample doesn't respond to our invitation to participate
  - Related to self-selection bias
- May be caused by errors related to **contact** or **cooperation** (Groves & Couper, 1998)
  - Contact = ability to find sample
  - Cooperation = ability to convince sample to respond
- Commonly measured through response rates
- But low response rates do not equal nonresponse bias (necessarily)



Nonresponse bias

- A standard mathematical definition of nonresponse bias is:

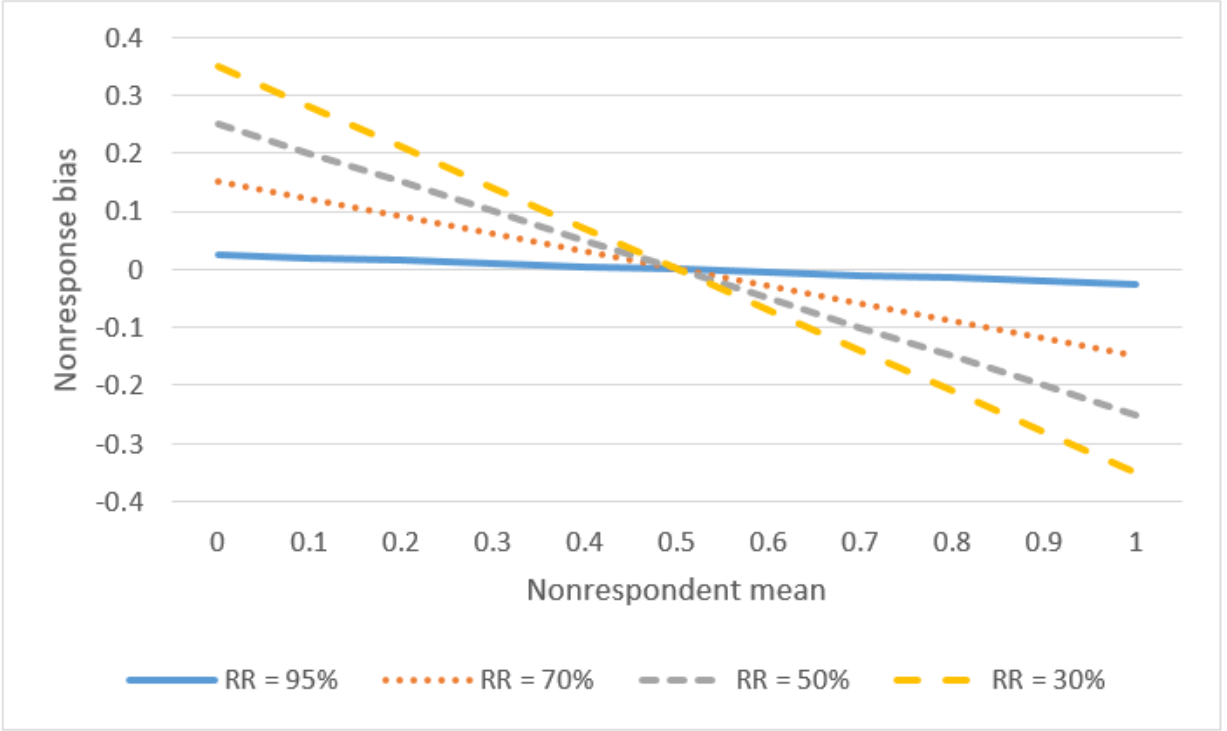
$$Bias(\bar{y}_r) = (1 - RR)(\bar{Y}_r - \bar{Y}_{nr})$$

$Bias(\bar{y}_r)$ = bias of respondent mean
$\bar{y}_r$ = respondent mean
RR = response rate
$\bar{Y}_r$ = respondent population mean
$\bar{Y}_{nr}$= nonrespondent population mean

What has more of an effect on bias: the response rate or how different the nonrespondents are from the respondents?

Nonresponse bias example (Groves & Couper, 1998)

- Assume $\bar{y}_r$ = 0.5
- At a high RR (e.g., 95%), big differences between $\bar{y}_r$ and $\bar{y}_{nr}$ result in small nonresponse bias
- If $\bar{y}_r = \bar{y}_{nr}$ then response rate doesn't matter as much
- If response rates are out of our control, we need to work to minimize the difference between $\bar{y}_r$ and $\bar{y}_{nr}$ to reduce bias
- Remember: bias is variable dependent



Response rates and representation

- Multiple examples that show the correlation between response rates and nonresponse bias is often weak
  e.g., Keeter et al. (2000); Curtin et al. (2000); Groves (2006); Groves and Peytcheva (2008)
- Response rate needs to be increased substantially to lower the average nonresponse bias (Brick & Tourangeau, 2017)
- Response rates, while important, are only one part of the equation
  - Increasing response rates is common prescription for surveys
- Saying that a survey with a low response rate (e.g., 50%, 25%, 10%) is not “adequate” can be misleading and undermines the work we are trying to accomplish

Response rates of population surveys

Survey	Year	Sample Size (Target)	Population	Response Rate	Sampling frame	Mode
California Health Interview Survey (CHIS)	2019	20,000	California	10.8% (Adult)	Stratified address-based sample	Web and telephone
California Behavioral Risk Factor Surveillance Survey (CA-BRFSS)	2019	10,000	California	40.2%	Random-digit dialing	Telephone
National Health Interview Survey (NHIS)	2018	30,000	United States	64.2% (Household)	Multistage, cluster sample	Face-to-face
National Health and Nutrition Examination Survey (NHANES)	2017-2018	5,000	United States	51.9% (Interview) 48.9% (Examination)	Multistage, cluster sample	Face-to-face

Note. Pew Research Center, Gallup, and other major surveys report average response rates consistently below 10% in recent years, but still considered trusted sources (Lavrakas et al., 2017; Marken, 2018; Kennedy & Hartig, 2019).

National Health and Nutrition Examination Survey (NHANES)

- Multistage, cluster sample of the United States
- 5,000 persons per year
- Face-to-face interview followed by an examination at a mobile examination center (MEC)
- National sponsorship
- Large incentives ($100+) for interview and examinations

NHANES	Interview RR	Examination RR
1999-2000	82.0	76.0
2001-2002	84.0	80.0
2003-2004	79.0	76.0
2005-2006	80.5	77.4
2007-2008	78.4	75.4
2009-2010	79.4	77.3
2011-2012	72.6	69.5
2013-2014	71.0	68.5
2015-2016	61.3	58.7
2017-2018	51.9	48.8

California Health Interview Survey (CHIS)

- Stratified sample of Californians
- 20,000 adult interviews per year
- As of 2019, design is a mixed-mode survey (web and telephone) from an address-based sample (ABS)
  - Before 2019, random-digit dial (RDD) telephone interview
- University sponsorship
- $2 pre-incentive

CHIS	Screener RR†	Adult RR
2001	59.2	37.7
2003	55.9	33.5
2005	49.8	26.9
2007	35.5	18.3
2009	34.9*	15.6
2011-2012	31.8*	16.5
2013-2014	29.0*	15.1
2015-2016	20.1*	9.1
2017-2018	8.0	3.6
2019**	14.9	10.8

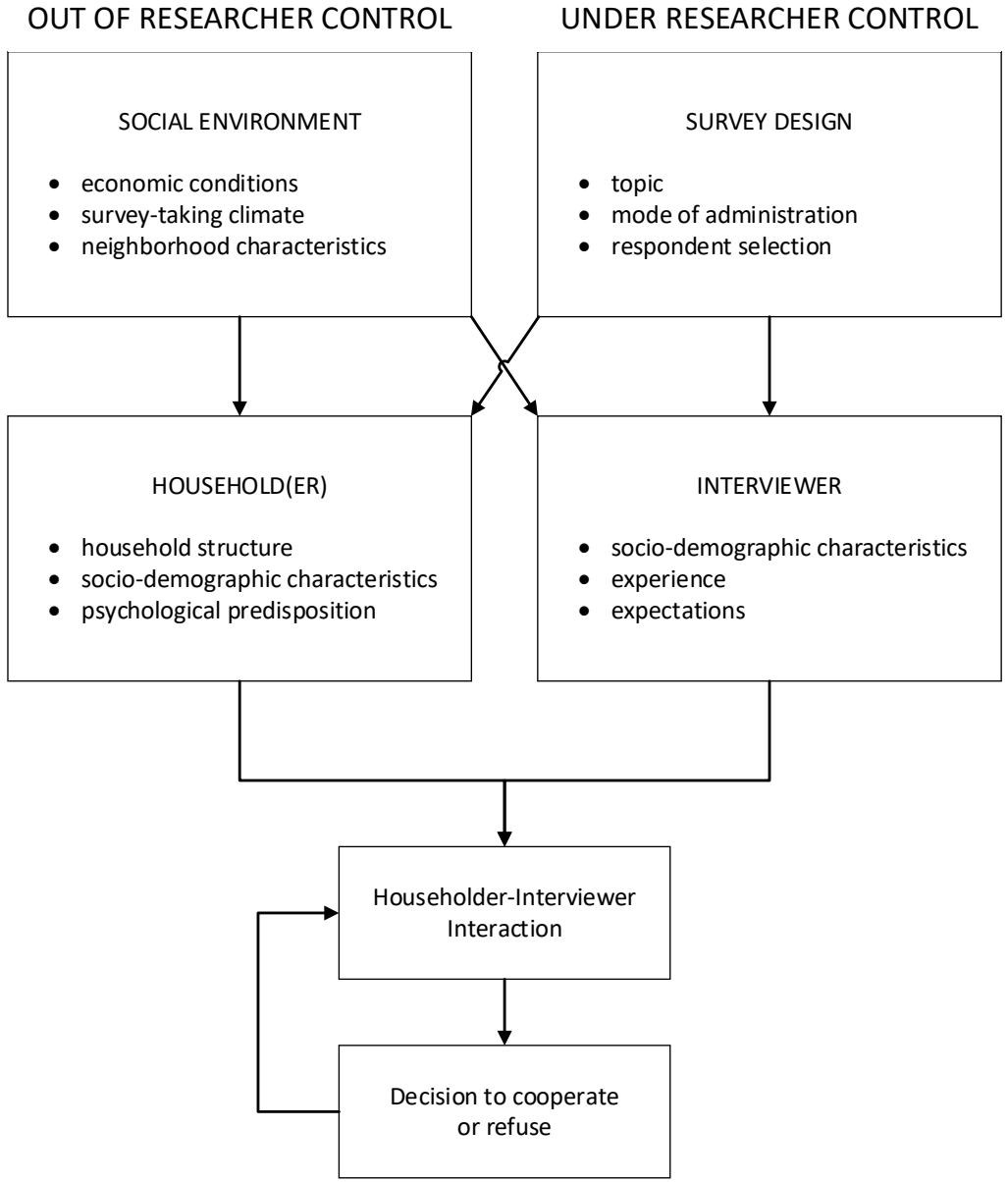
† Respondents who completed the screener portion of the survey verifying study eligibility
* Estimated due to separate reported landline and cell screener response rates
** Represents a methodological redesign from RDD/CATI to ABS/mixed-mode

Comparing response rates

- American Association for Public Opinion Research (AAPOR) Standard Definitions is an international standard for calculating and reporting response rates (AAPOR, 2016)
- Inappropriate to compare response rates of dramatically different study designs
  - Apples : Oranges :: NHANES : CHIS
- Different study attributes (e.g., incentives, topic, survey mode) influence cooperation success

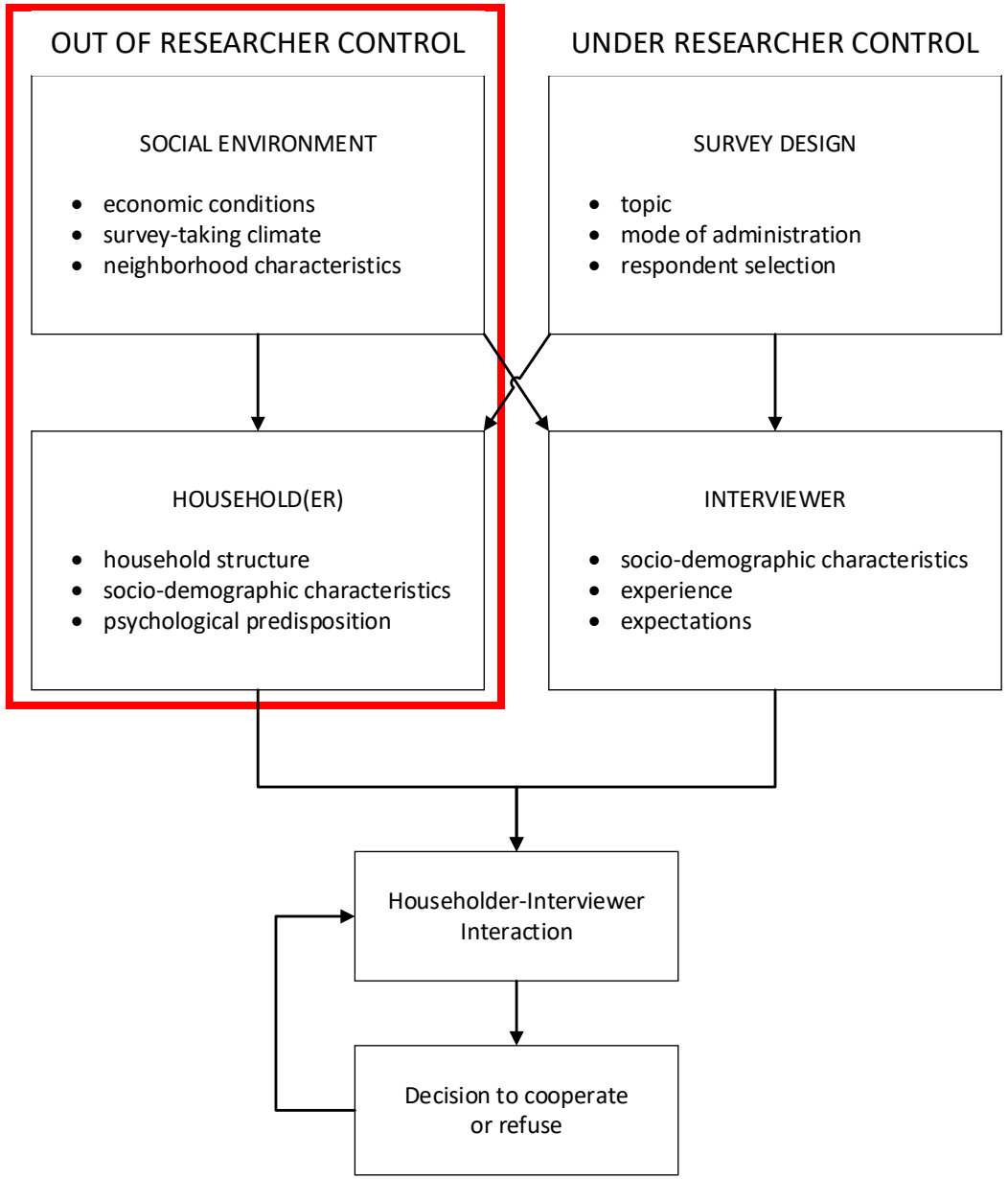
Survey cooperation

- Conceptual framework for survey cooperation
  (Groves & Couper, 1998)
- Framework is used primarily for face-to-face surveys, but the general framework still applies to other designs



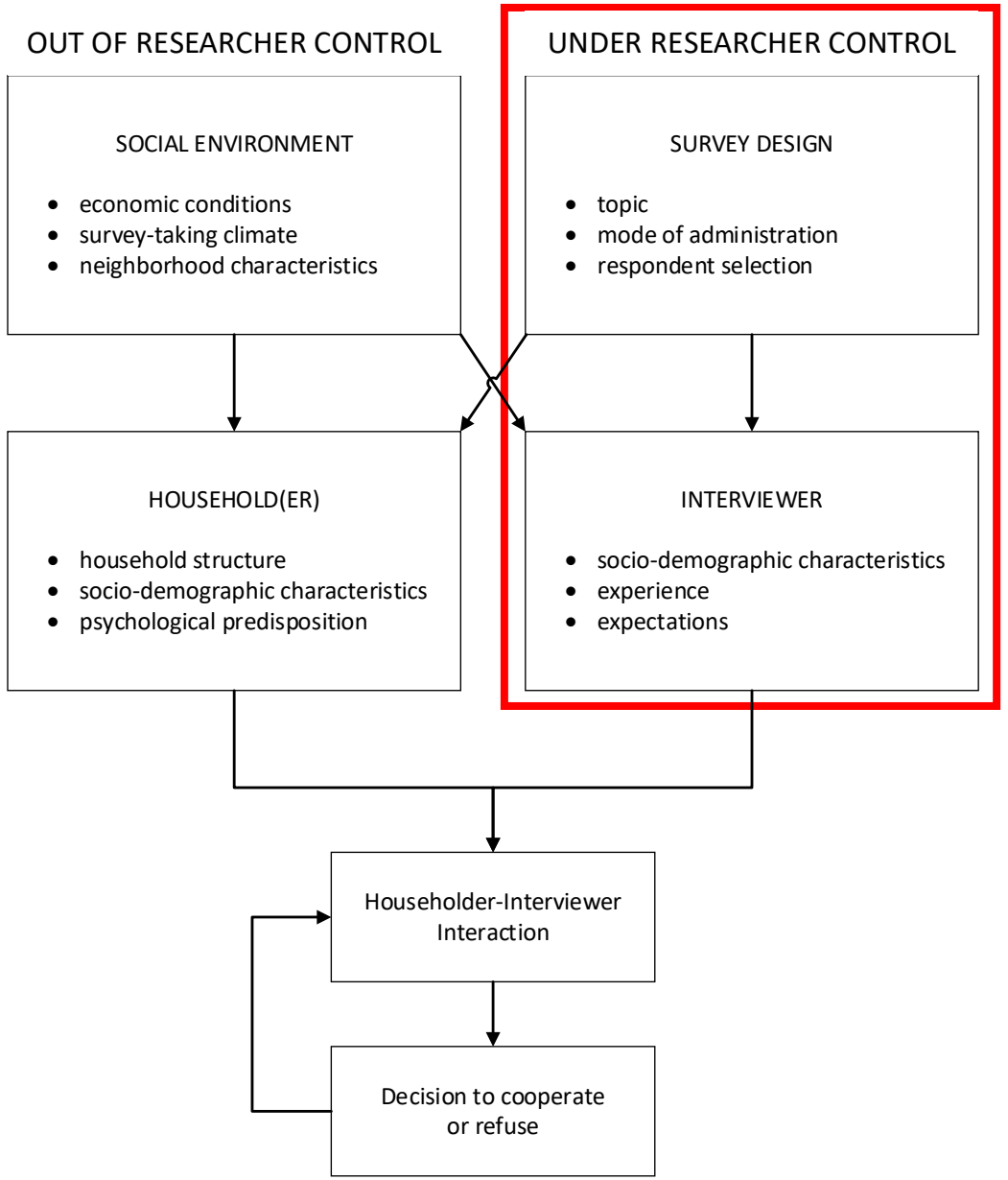
Out of (our) control...

- Social Environment
  - Survey-taking climate
    - Declining response rates, reluctance
    - Distrust, discrediting claims
    - COVID-19 restrictions, fear
  - Neighborhood characteristics
    - Areas disproportionately effected
- Household(er)
  - Socio-demographic characteristics
  - Psychological predisposition



Under (our) control…

- Survey design
  - Topic
  - Mode of administration
  - Respondent selection
  - Incentive
- Interviewer (when applicable)
  - Or other study staff (e.g., nurses)
  - Or the instrument interface (e.g., paper questionnaire, web survey)



Who often doesn’t respond to surveys

- Young adults (age 18-29)
- Low income/social economic status (SES) households
- Limited English proficient speakers
- Less socially-connected or politically active
- Racial/ethnic minorities

- But it depends on the survey, and each survey is different...

A note on weighting

- A key feature of probability samples is that we can weight the final sample to be representative of the population
  - Need known population totals (or at least totals you are confident in)
    - For example, California Department of Finance (DOF) population estimates, U.S. Census, American Community Survey (ACS)
    - Might include age, gender, race, ethnicity
- Weighting reduces bias due to nonresponse, sampling, and coverage
  - Reductions maximized if the weighting variables are correlated with both response to the survey and the survey outcome (Little & Vartivarian, 2005)
- Two different samples can produce similar estimates when weighted to a population using equivalent methods
- But weighting is not a silver bullet or magic wand

California Health Interview Survey weighting

- CHIS weights to population totals in both ACS and California DOF
- CHIS weights to various combinations of
  - Age
  - Gender
  - Race/ethnicity
  - Education
  - County and region
  - Housing tenure
  - Number of adults in household

Weighting in population-based surveys that collect biomeasures

- Many population-based surveys use sample weights to estimate biomarker outcomes

- National Health and Nutrition Examination Survey (NHANES)
  - https://wwwn.cdc.gov/nchs/nhanes/tutorials/module3.aspx
- Health and Retirement Study (HRS)
  - Ofstedal et al. (2011)
- University of Michigan Dioxin Exposure Study (UMDES)
  - Garabrant et al. (2009)

Trade-Offs of the Possible (T.O.P.)

- Types of biomeasures
- Sample design
- Mode of data collection
- Incentives
- Location and travel

T.O.P. : Types of biomeasures

Benefits

- Minimally invasive methods allow for
  - More survey mode flexibility
  - Reduction in costs for study staff (e.g., nurses) and travel

Limitations

- Minimally invasive methods may not allow for collection of specific biomarkers of interest

T.O.P. : Sample design

Benefits

- A population-based sample design allows selection of phone numbers or addresses with known probabilities
- Can follow-on to existing design to reduce screening, improve individual level data

Limitations

- Concerns over low sample sizes, poor response rates and representation
- Suffer limitations and errors of the original design

T.O.P. : Sample design (cont.)

Benefits

- Non-probability designs or convenience samples can be easier to implement

Limitations

- We are less confident in non-probability design's variability, impact of self-selection bias

T.O.P. : Mode of data collection

Benefits

- Web survey can be relatively inexpensive to collect basic survey information

Limitations

- Mail, web, or telephone require:
  - Change biomeasure collection, or
  - Two-stage contact to allow for face-to-face testing

T.O.P. : Incentives

Benefits

- Incentives (of any amount) increase response rates
- Pre-incentives can be more effective than promised incentives

Limitations

- Have to balance final sample size target with budget for incentives
- Pre-incentives are a big upfront cost with no guarantee of return

T.O.P. : Location and travel

Benefits

- Convenient locations or guaranteed transportation increases cooperation
- In-home visits reduce respondent burden

Limitations

- Too many locations spreads the research team too thin, too few locations is inconvenient
- May not be able to sufficiently compensate for respondent travel
- In-home visits costly when using nurses or phlebotomists, difficulty taking necessary equipment

Conclusion

- We can’t let perfect be the enemy of good
- Low response rates are not desirable, but doesn’t mean that the data is of no value or “inadequate”
- Accept that our design will have errors, but that we can work to minimize the errors that we have (some) control over
- Adopting principles of good survey design for materials and recruitment and the use of population-based methods may go far in improving the usability what we can collect

Additional resource

- From the CDC: Guidance, Examples and Tools for Probability Sampling when Designing a Population-based Biomontoring Study

Thank you!

Brian M. Wells, PhD

bmwells@ucla.edu

References

- American Association for Public Opinion Research. (2016). *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys.* 9th Edition. Chicago, IL: AAPOR.
- Biemer, P. P., & Lyberg, L. E. (2003). *Introduction to survey quality*. John Wiley & Sons.
- Brick, J. M., & Tourangeau, R. (2017). Responsive survey designs for reducing nonresponse bias. *Journal of Official Statistics*, *33*(3), 735-752.
- California Health Interview Survey. (2020). *CHIS 2019 Methodology Series: Report 5 - Weighting and Variance Estimation*. Los Angeles, CA: UCLA Center for Health Policy Research.
- Centers for Disease Control and Prevention (2020). *Guidance, Examples and Tools for Probability Sampling when Designing a Population-based Biomontoring Study*. Retrieved from https://www.cdc.gov/nceh/tracking/pdfs/BiomonitoringStudyGuidance-508.pdf
- Curtin, R., Presser, S., & Singer, E. (2000). The effects of response rate changes on the index of consumer sentiment. *Public Opinion Quarterly, 64*(4), 413-428.
- Garabrant et al. (2009). The University of Michigan Dioxin Exposure Study: Predictors of Human Serum Dioxin Concentrations in Midland and Saginaw, Michigan. *Environmental Health Perspectives, 117*(5), 818-824.
- Groves, R. (2006). Nonresponse rates and nonresponse bias in household surveys. *Public Opinion Quarterly, 70*(5), 646-675.
- Groves, R. M. & Couper, M. P. (1998). *Nonresponse in household interview surveys.* New York: John Wiley & Sons.
- Groves, R. and Peytcheva, E. (2008). The impact of nonresponse rates on nonresponse bias. *Public Opinion Quarterly, 72*(2), 167-189.

References (continued)

- Jaszczak, A., Lundeen, K., & Smith, S. (2009). Using nonmedically trained interviewers to collect biomeasure in a national in-home survey. *Field Methods, 21*(1), 26-48.
- Keeter, S., Miller, C., Kohut, A., Groves, R., & Presser, S. (2000). Consequences of reducing nonresponse in a national telephone survey. *Public Opinion Quarterly, 64*(2), 125-148.
- Lavrakas, P, et al. (2017). The future of U.S. general population telephone survey research. Retrieved from https://www.aapor.org/Education-Resources/Reports/The-Future-Of-U-S-General-Population-Telephone-Sur.aspx
- Lindau, S. T. & McDade, T. W. (2007). Minimally invasive and innovative methods for biomeasure collection in population-based research. In M. Weinstein, J. W. Vaupel, & K. W. Wachter (Eds.). *Biosocial surveys* (pp. 251-277). National Academies Press.
- Little, R. J. & Vartivarian, S. (2005). Does weighting for nonresponse increase the variance of survey means? *Survey Methodology, 31*(2), 161-168.
- Ofstedal, M. B., Weir, D. R., Chen, K. T. J., & Wagner, J. (2011). Updates to HRS sample weights. *HRS documentation report, 13.* Ann Arbor, MI: University of Michigan.
- Sakshaug, J. W., Ofstedal, M. B., Guyer, H., & Beebe, T. J. (2015). The collection of biospecimens in health surveys. In T. P. Johnson (Ed.), *Handbook of Health Survey Methods* (pp. 383-419). Hoboken, New Jersey: John Wiley & Sons, Inc.
- Weir, D. (2008). Elastic powers: The integration of biomarkers into the Health and Retirement Study. In M. Weinstein, J.W. Vaupel, & K. Wachter (Eds.), *Bio-Social Surveys: Current insight and future promise* (pp. 78-95). Washington DC: The National Academies Press.