Validating the results of a nationally representative probability-based panel survey for COVID-19, United States, 2020–2022

Divya Vohra, Holly Matulewicz, Andy Weiss, Willow Crawford-Crudell, Chandra Couzens, Nancy Clusen

Federal Committee on Statistical Methodology Conference 2023

Session A-3: Perspectives from the COVID-19 Pandemic

Tuesday, October 24, 2023

Acknowledgements

/ **U.S. Centers for Disease Control and Prevention**

- Funding and support (no. RFA-DR-21-087.2)

/ **Ipsos**

/ **Coauthors at Mathematica**

Survey background and purpose

Public health emergency called for timely data to fill information gaps

- Surveillance data was not complete
- There was limited information on behaviors of cases, contacts after notification
- Landscape in pandemic could change quickly

Robert Santos (2014) fit-for-purpose framework balances:

- available resources
- rigor of research design and implementation
- nature of the insights needed to effectively address the research questions

Key concerns about data quality from panel

- **Bias from panel methodology: high completion rate, but low overall response rate**
- **Self-reported data: recall error from long reporting period**

Survey design and administration

- **Ipsos Knowledge Panel**
  - Probability-based sample, representative of adult U.S. population
  - Self-administered, voluntary online survey
  - Available in English and Spanish
- **Pre-test interviews**
- **Survey fielded Feb-March 2022**
  - 4-week field period

- **22,514 panel members sampled**
- **15,923 offered eligibility screener completed (70%)**
- **9,269 met eligibility criteria and completed (58%):**
  - Positive SARS-CoV-2 test result (cases, n: 9,269), or
  - Notification of exposure (contacts, n: 5,369)
- **Response rate 4%** (AAPOR 2023)
- **Weighted to CPS, ACS**

Research questions

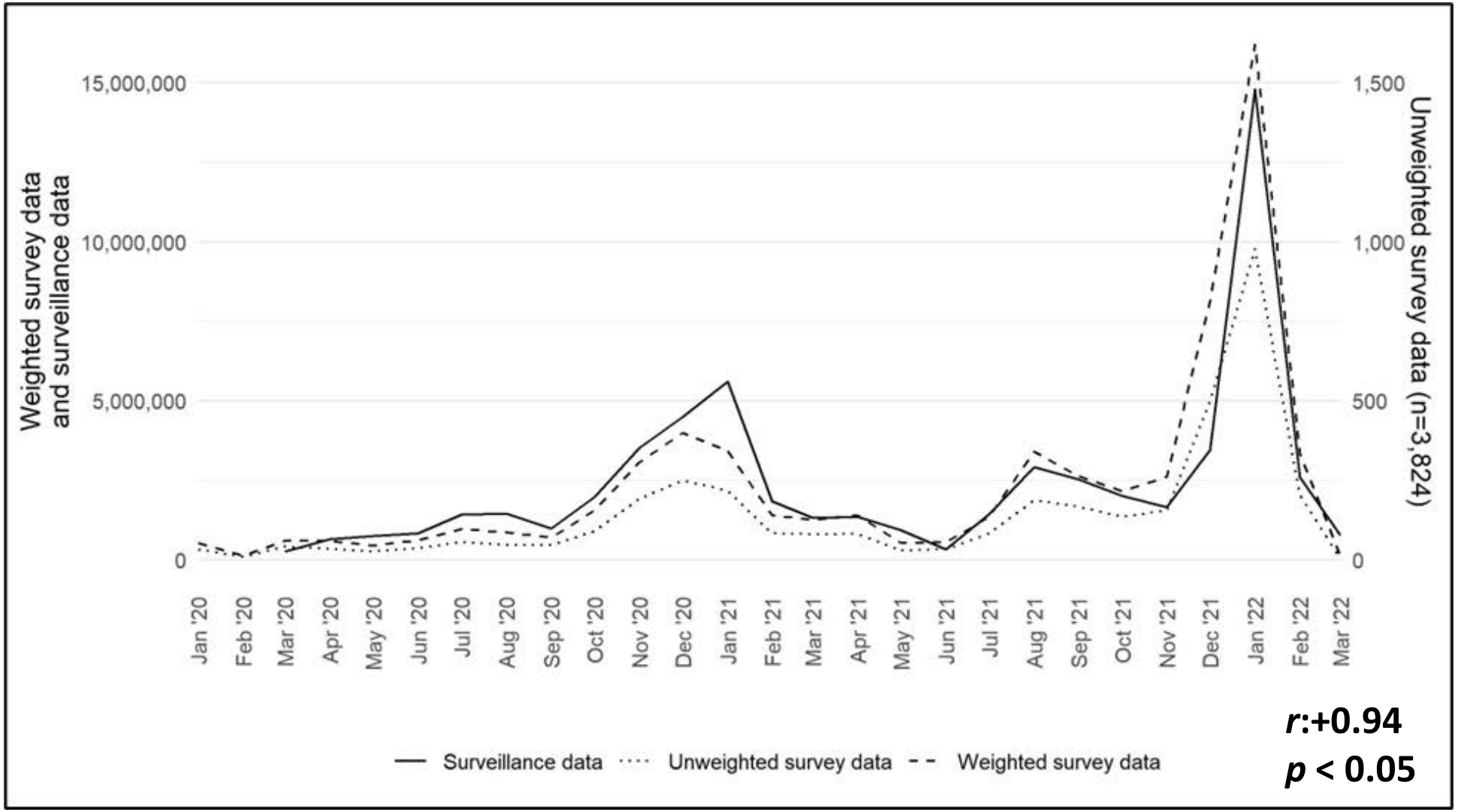
1. Overall, how well did the case-based survey data align with CDC surveillance data of the number of reports of all adults (aged 18 years or older) who tested positive for SARS-CoV-2?

2. How well did the case-based survey data align with CDC surveillance data of the number of reports of all adults who tested positive for SARS-CoV-2 by select demographic characteristics?

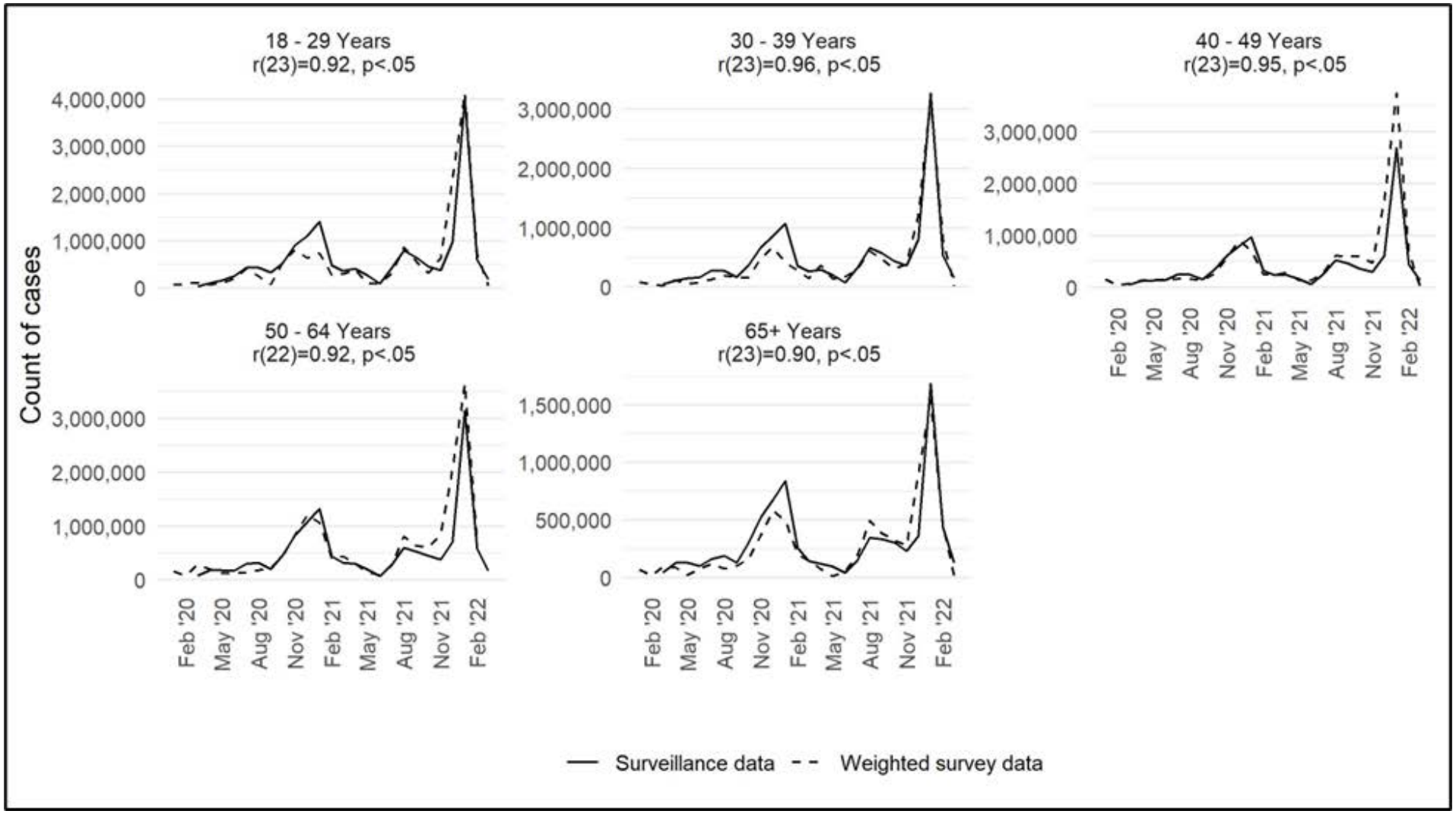
Methods used to answer these questions

- **Compared survey results to CDC surveillance data in the same time period**
  - March 2020 to March 2022
  - Subtracted fatalities from surveillance data
- **Created epidemiological curve**
  - Aggregate cases by month, year for both data sets
  - Visualize distribution of cases overall, then by age, sex, race/ethnicity
- **Calculated Pearson's correlation coefficients (r) and *p* values**
  - Overall and age, sex, race/ethnicity

Findings on Q1: Survey data highly correlated with CDC surveillance data on overall case counts by month



Findings on Q2: Survey data remained highly correlated with CDC surveillance data across age groups and other demographics



Findings in summary

- **Unique opportunity for comparison survey panel data with surveillance data**
  - Found high correlation overall and by age group
- **Findings validated that survey data mirrored U.S. adult population of cases**
  - Strengthens confidence in the data
    - Despite low response rate
    - Combined with pre-test results, allays concerns about reporting error for long recall period
- **Provides quality data to inform decision making**
  - Fills knowledge gaps on behaviors of cases, contacts
  - Opportunities to enhance findings from surveillance data using demographics, other covariates

Limitations in each data set

Panel survey data

- Self-administration data subject to reporting error
- Does not include:
  - Adults in group quarters
  - Deceased prior to survey period
  - Adults with language or literacy barriers

CDC surveillance data

- Cases identified and reported to CDC
- Missing values (race, ethnicity)

Key takeaways

1. **Low response rates do not always correlate with poor quality data**
   - Our respondents closely mirrored CDC surveillance data
   - Validation our results on case counts by month mirrored the available gold standard

2. **Panel survey met the need: fit-for-purpose data collection**
   - Timely, valid data
   - Low(er) cost, recruitment costs shared across studies
   - Filled critical information gaps about behaviors of cases and contacts (Oeltmann et al. 2023)

Contact us for more information

/ Holly Matulewicz, Principal Survey Researcher

- hmatulewicz@mathematica-mpr.com

/ Andy Weiss, Senior Fellow

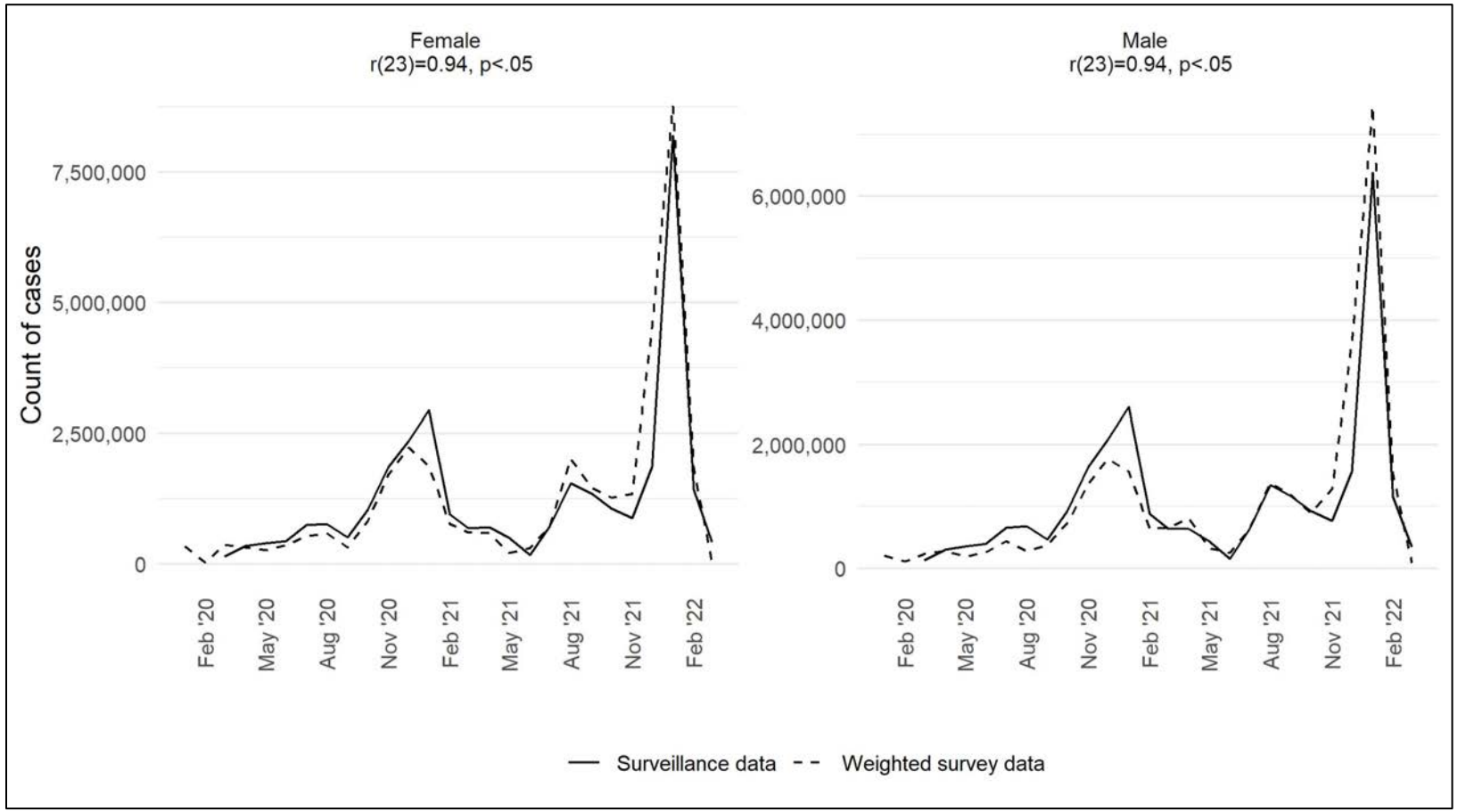
- aweiss@mathematica-mpr.com

Resources

- American Association for Public Opinion Research. 2023. “Standard Definitions. Final Dispositions of Case Codes and Outcome Rates for Surveys.” Available at https://aapor.org/standards-and-ethics/standard-definitions/.
- Centers for Disease Control and Prevention. 2023a. “COVID-19 Case Surveillance Public Use Data with Geography.” Available at https://data.cdc.gov/Case-Surveillance/COVID-19-Case-Surveillance-Public-Use-Data-with-Ge/n8mc-b4w4. (Accessed August 11, 2023).
- Centers for Disease Control and Prevention. 2023b. “COVID-19 Weekly Cases and Deaths by Age, Race/Ethnicity, and Sex.” Available at https://data.cdc.gov/Public-Health-Surveillance/COVID-19-Weekly-Cases-and-Deaths-by-Age-Race-Ethni/hrdz-jaxc. (Accessed August 11, 2023).
- Dever J.A., Amaya A., Srivastav A., Lu P.J., Roycroft J., Stanley M., Stringer M.C., Bostwick M. G., Greby S. M., Santibanez T.A., Williams W. W. 2021. “Fit for Purpose in Action: Design, Implementation, and Evaluation of the National Internet Flu Survey.” *Journal of Survey Statistics and Methodology* 9: 449–76.
- Groves R.M., Peytcheva E. 2008. “The Impact of Nonresponse Rates on Nonresponse Bias: A Meta-Analysis.” *Public Opinion Quarterly* 72:167–89.
- Hays R.D., Liu H., Kapteyn A. 2015. “Use of Internet Panels to Conduct Surveys.” *Behavior Research Methods* 47:685–90.
- Ipsos. “Public Affairs KnowledgePanel.” Available at https://www.ipsos.com/en-us/solutions/public-affairs/knowledgepanel (accessed May 5, 2023).
- Oeltmann J.E., Vohra D., Matulewicz H.H., DeLuca N., Smith J.P., Couzens C., Lash R.R., Harvey B., Boyette M., Edwards A., Talboy P.M., Dubose O., Regan P., Loosier P., Caruso E., Katz D.J., Taylor M.M., Moonan P.K. Isolation and Quarantine for Coronavirus Disease 2019 in the United States, 2020-2022. Clin Infect Dis. 2023 Jul 26;77(2):212-219.
- Santos R. L. 2014. “Presidential Address: Borne of a Renaissance—A Metamorphosis for Our Future.” *Public Opinion Quarterly* 78: 769–77.

Questions?

Supplemental Table 1. Monthly COVID-19 case counts in the United States by sex and data source (CDC surveillance and weighted survey), excluding fatalities, January 2020 to March 2022



Supplemental table 2. Monthly COVID-19 case counts in the United States by race/ethnicity and data source (CDC surveillance and weighted survey), excluding fatalities, January 2020 to March 2022