



# Paradata Analysis of Participant Recruitment and Retention for COVID Household Transmission Study

**Xiaoshu Zhu, PhD | Steph Battan-Wraith, PhD |  
Vanessa Olivo, MPH | Kerry Grace Morrissey, MPH**

WESTAT @ FCSM 2024

The views presented are those of the author(s) and do not represent the views of any Government Agency/Department or Westat



# Background



# Respiratory Virus Transmission Network – National

- Purpose

- Estimate the transmission or secondary attack rate of COVID-19 within a household
- Estimate the effectiveness of vaccination in a national population

- Study Design

- Achieve greater geographic diversity and inclusion in the study population
- Case-ascertained virtual household transmission design



# Respiratory Virus Transmission Network – National

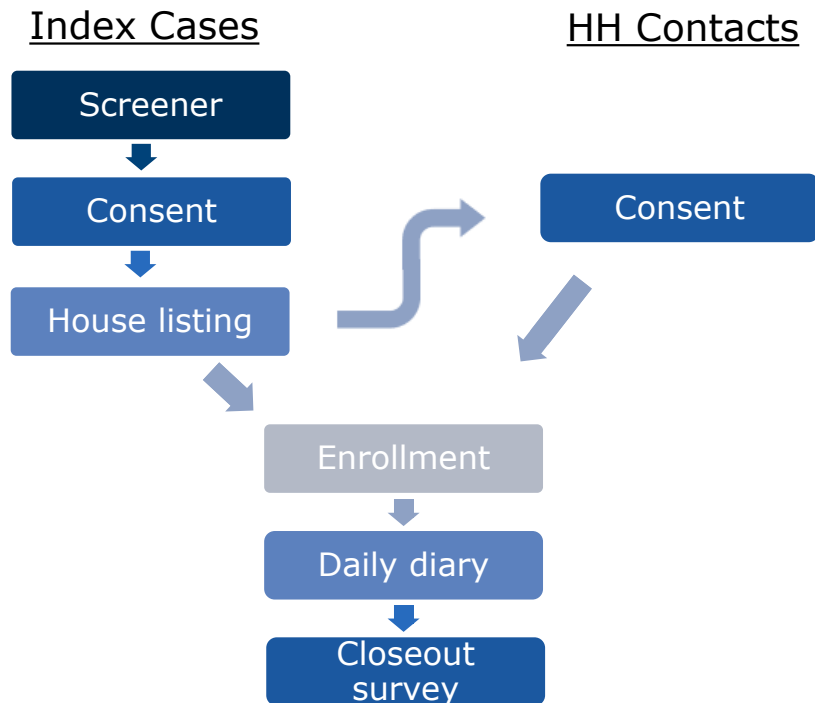
## ■ Key Features

- Web-based recruitment and enrollment of participant households
- Web-based data collection and at-home self-collected biospecimen collection
- Partnership with commercial entities

## At-home data and specimen collection



# Study Flow



## ▪ Index case

- Identified individuals with acute COVID infection
- More study components








## ▪ HH Contact


- Invited after index case consents to participate
- Fewer study components

# Participant Dashboard


**My Tasks**

[My Account](#) [View Track COVID at Home! Research Test Results](#)

 Enrollment Questionnaire	<a href="#">Go to Enrollment Questionnaire→</a>
 Diary Day - 2	Due: <b>5/6/2022</b>
 Diary Day - 3	Due: <b>5/7/2022</b>
 Diary Day - 4	Due: <b>5/8/2022</b>
 Diary Day - 5	Due: <b>5/9/2022</b>
 Diary Day - 6	Due: <b>5/10/2022</b>
 Diary Day - 7	Due: <b>5/11/2022</b>

 \$50

Receive when you complete enrollment, **collect and send COVID-19 tests, answer daily questions for 7 days and collect day 1 blood sample**

 \$25

After finishing the first 7 days, you may receive when you complete **3 more days of COVID-19 tests and daily questions** (in other words, when you complete all study activities for 10 days).

[Contact](#)

# Paradata

- Data generated as a by-product of the data collection process (Census, <https://www.census.gov/topics/research/paradata.html>)
- Why paradata?
  - Inform and improve questionnaire design
  - Monitor data collection process, evaluate data quality and cost
  - Assess respondent behavior and predict response propensity
- Current study
  - Change answers
  - Breakoff
  - Login attempts
  - Duration
  - Contact attempts
  - Device used

# Research Questions





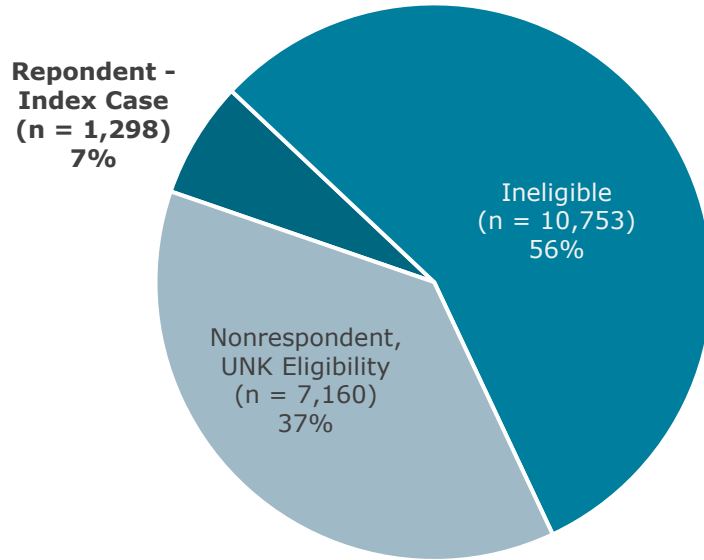
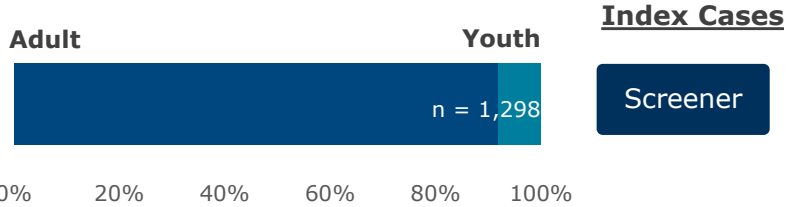
# Research Questions

- How was the burden placed on respondents by study components?
- How were contact attempts associated with completion status?
- How did the survey experience from previous study components affect participation status for the next step?

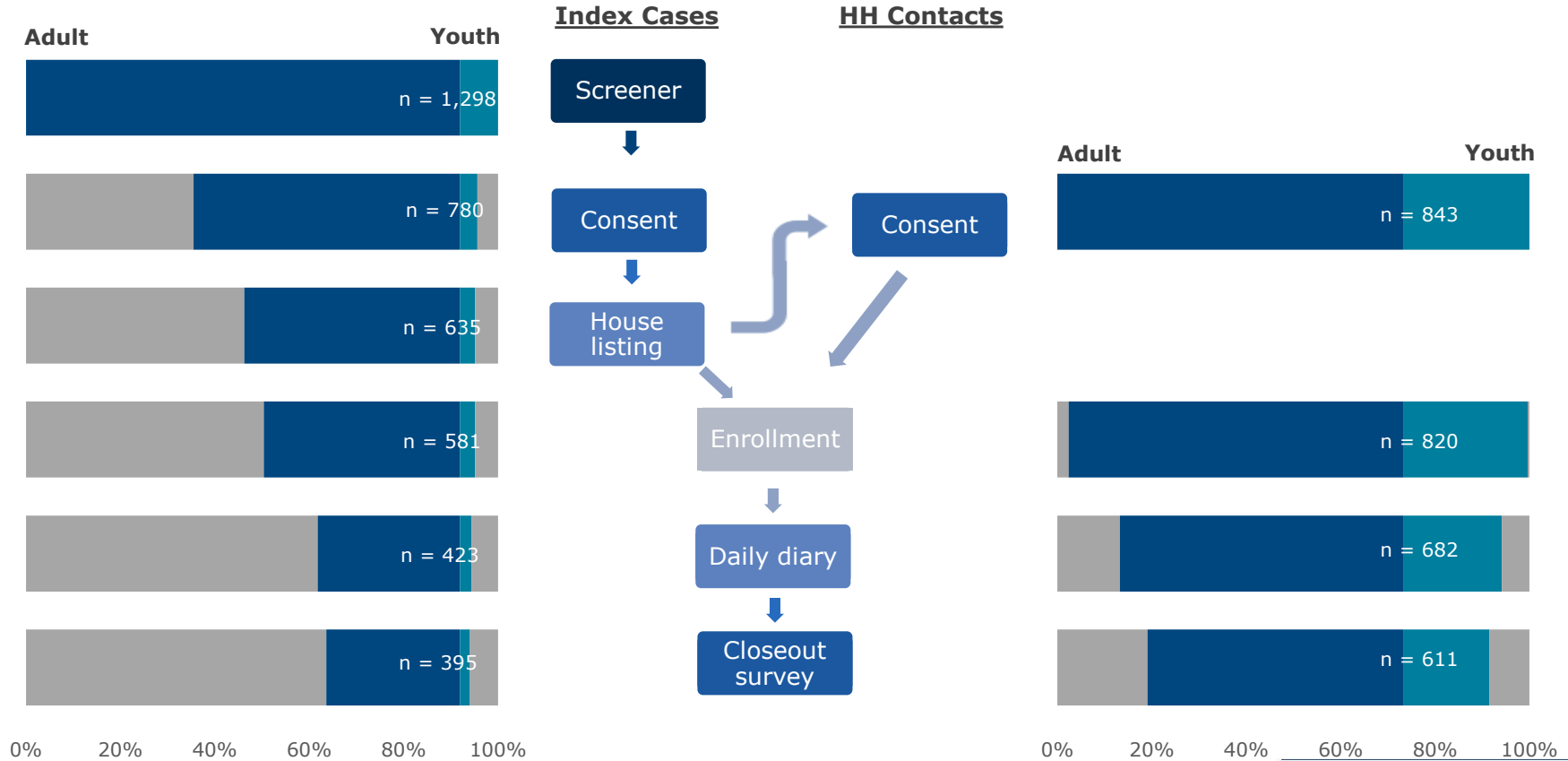
# Results



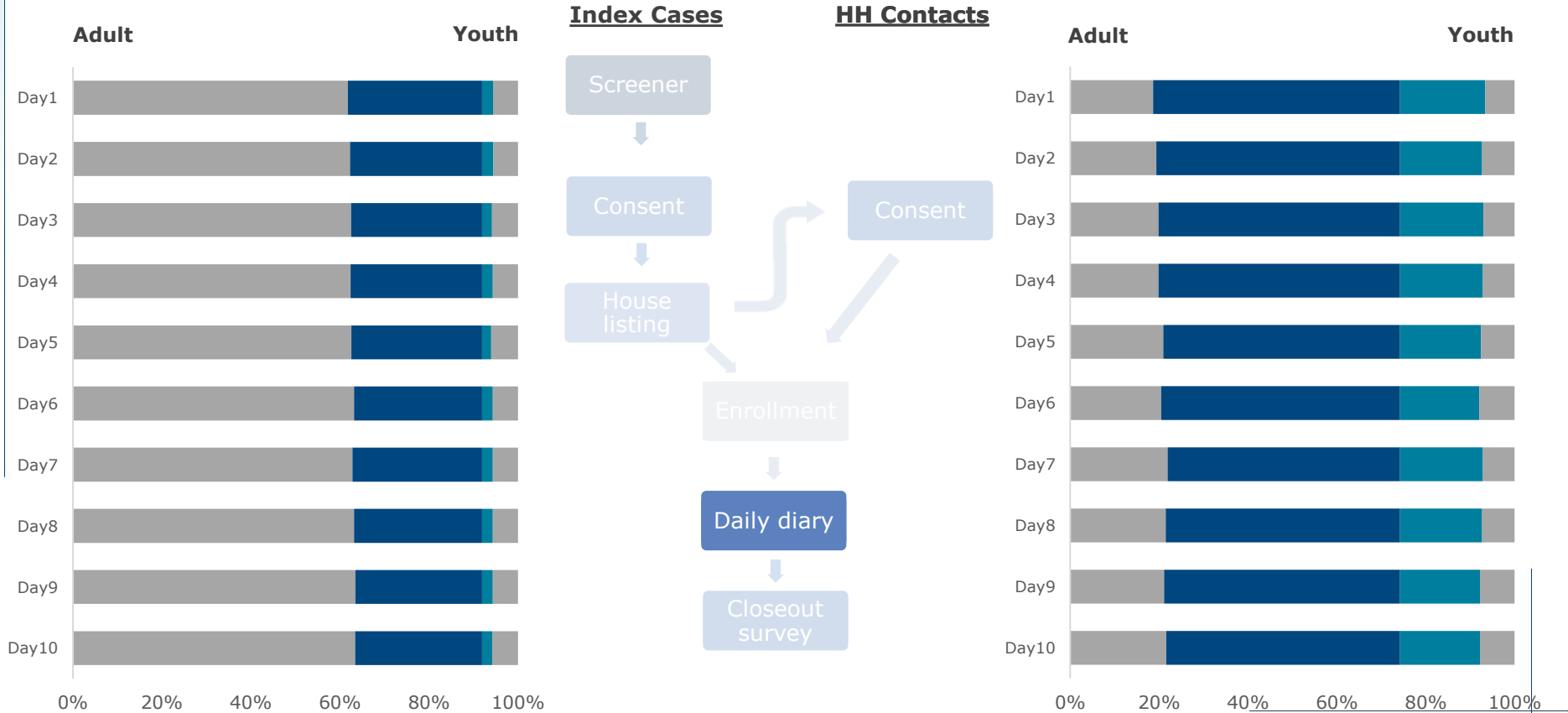
# Recruitment and Retention



# Recruitment and Retention (cont.)



# Recruitment and Retention (cont.) – Daily Diary



# Respondent Burden

## Index Cases

Screener



Consent



House listing

Enrollment



Daily diary



Closeout survey

## HH Contacts

Consent



### ▪ Change answer

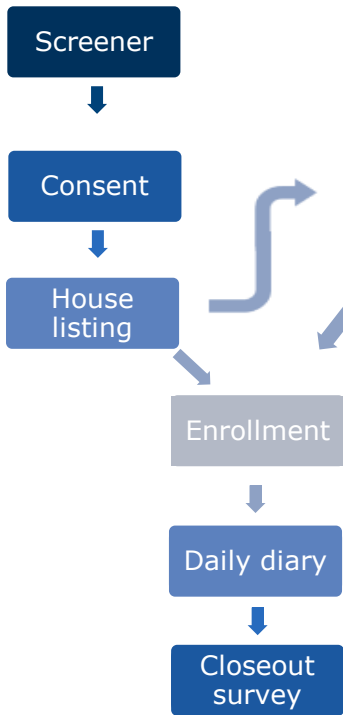
- Majority never changed answers
- Enrollment: Height, upload copy of vaccine card
- Daily diary: Collect a COVID test sample today, enter the COVID sample kit ID

### ▪ Page at breakoff

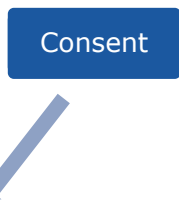
- No page with high frequency at breakoff among incompletes
- Screening: COVID Test date, Name, Relationship with index cases
- Enrollment: Lot number for the first COVID vaccine dose, upload copy of vaccine card

# Respondent Burden (cont.)

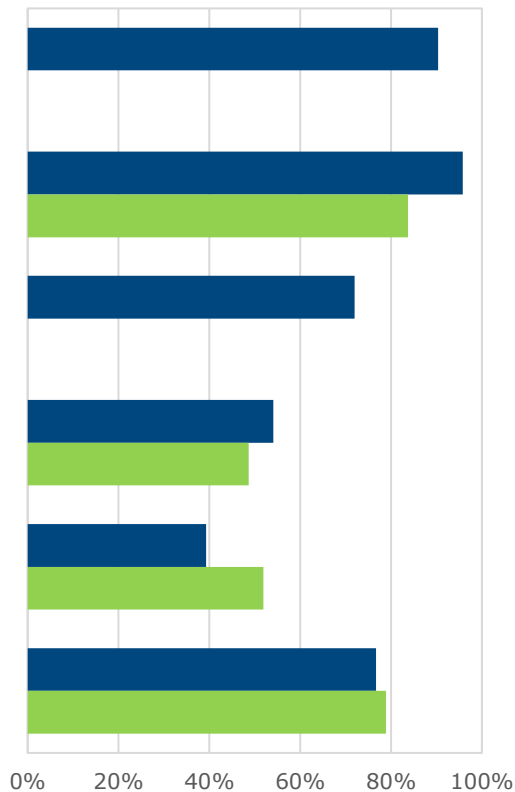
## Index Cases



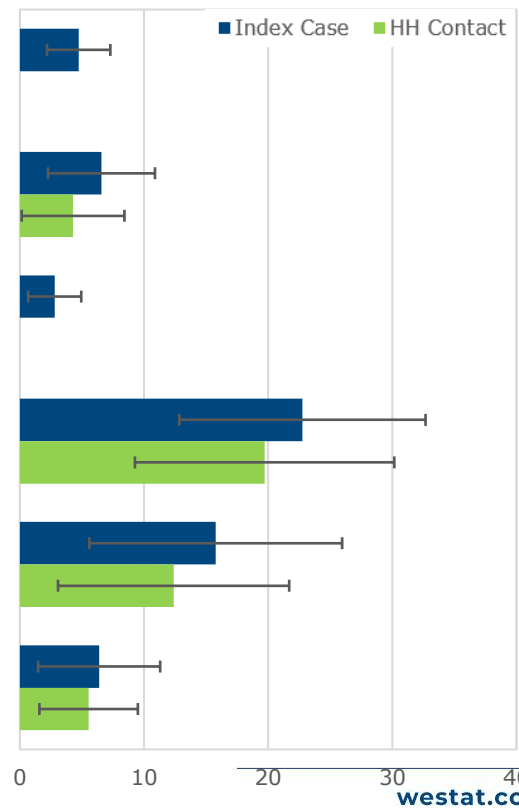
## HH Contacts



## Percent of One Login

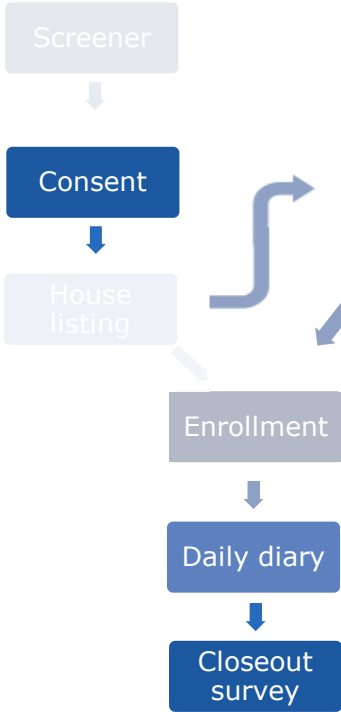


## Component Duration (mins)

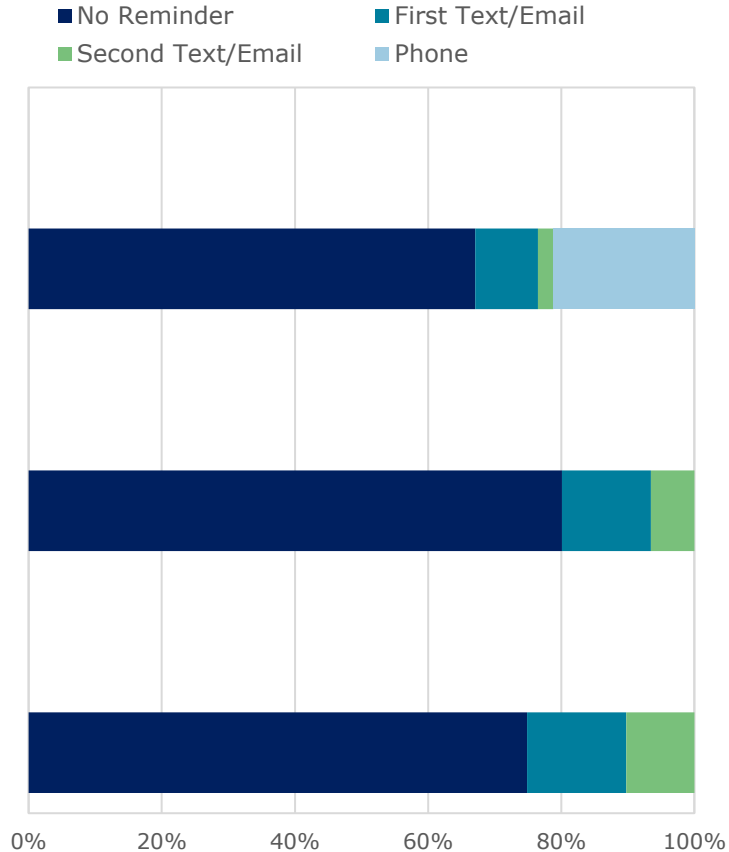
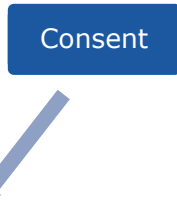


# Contact Attempts

## Index Cases



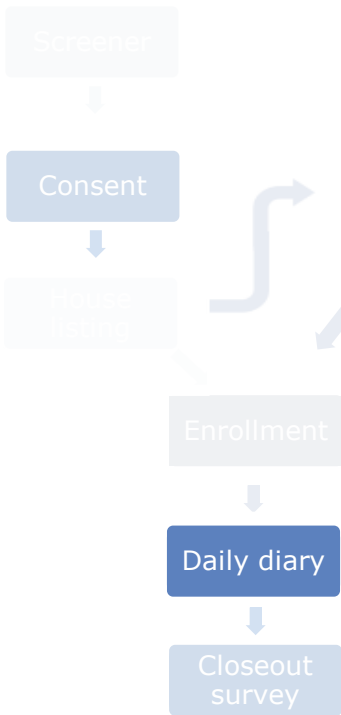
## HH Contacts



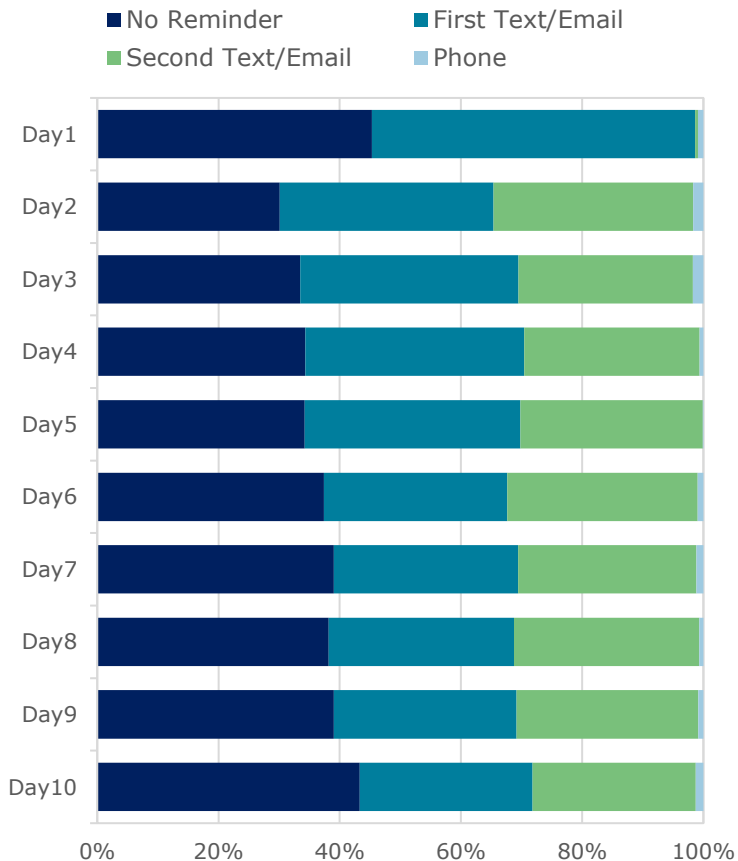


# Contact Attempts – Daily Diary

## Index Cases



## HH Contacts



# Survey Experience Affects Participation Status

## Importance rank in predicting next response

- Random forest and logistic regression

	Paradata				Person-level						HH-level			
	n. of Sess.	Sess. Durat.	Device	Contct Atmpt	Partic. Type	Proxy Resp.	Parent Permit	Age	Ethnic.	Race	HH Size	HH Type	Region	Recrt. Sourc.
Screen-> Consent		*									*			
Consent-> Enroll			*		*		*				*	*		
Enroll-> Diary		*			*		*				*	*		*
Diary -> Closeout			*	*	*		*	*		*	*		*	*

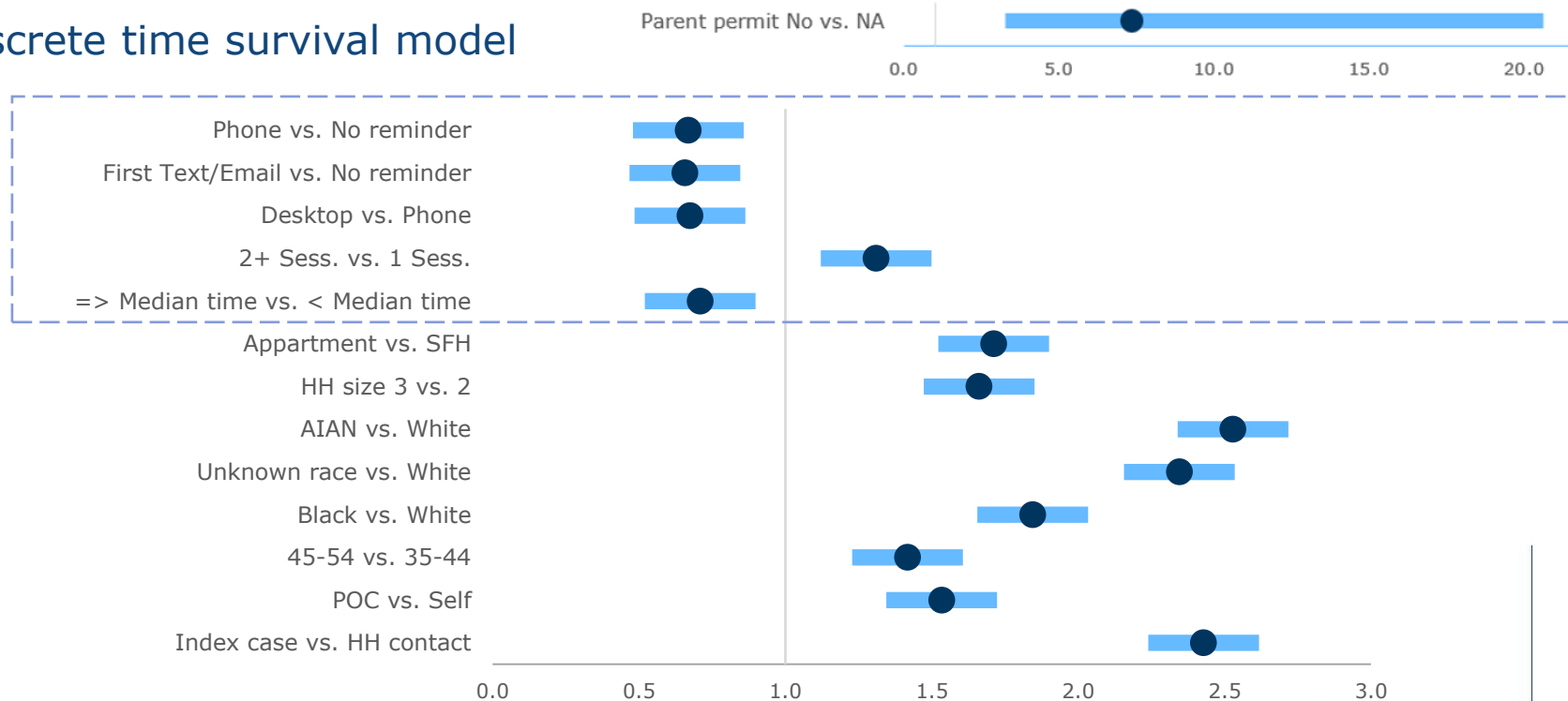
Note: \* means significant in logistic regression



# Survey Experience Affects Participation Status (cont.)

## Association with time to withdraw after consent

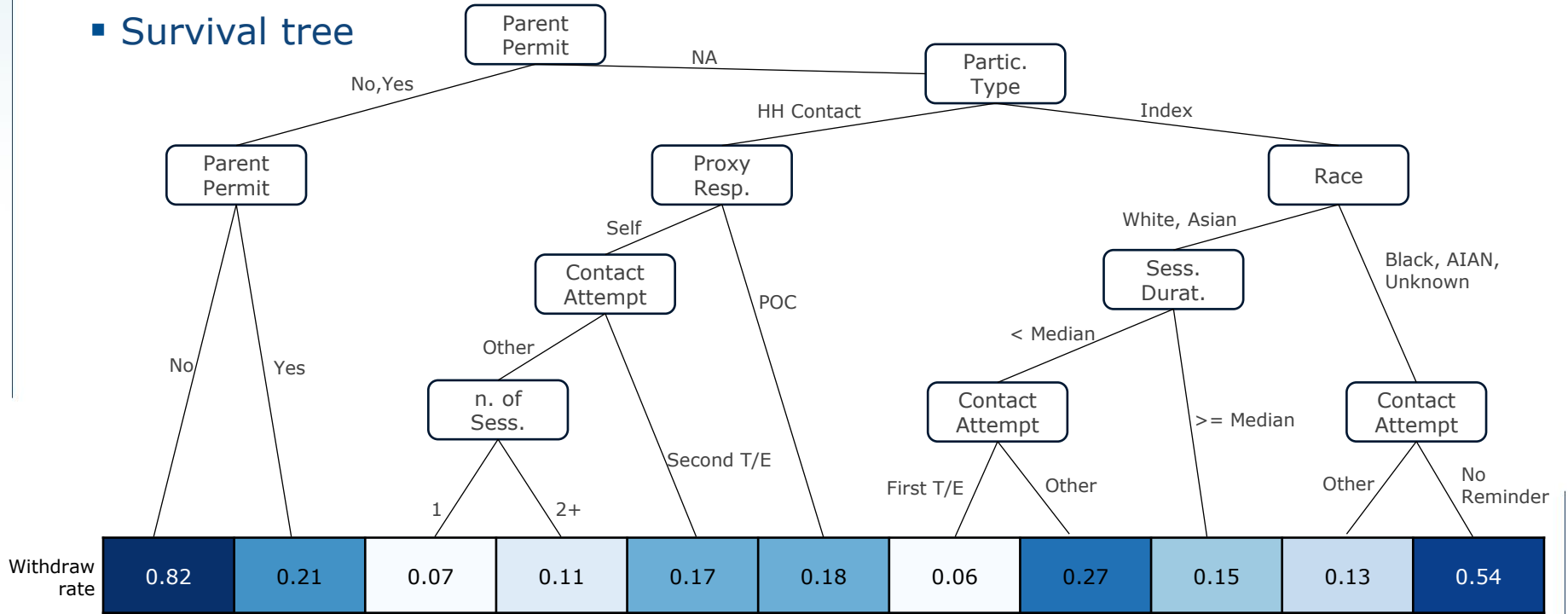
- Discrete time survival model



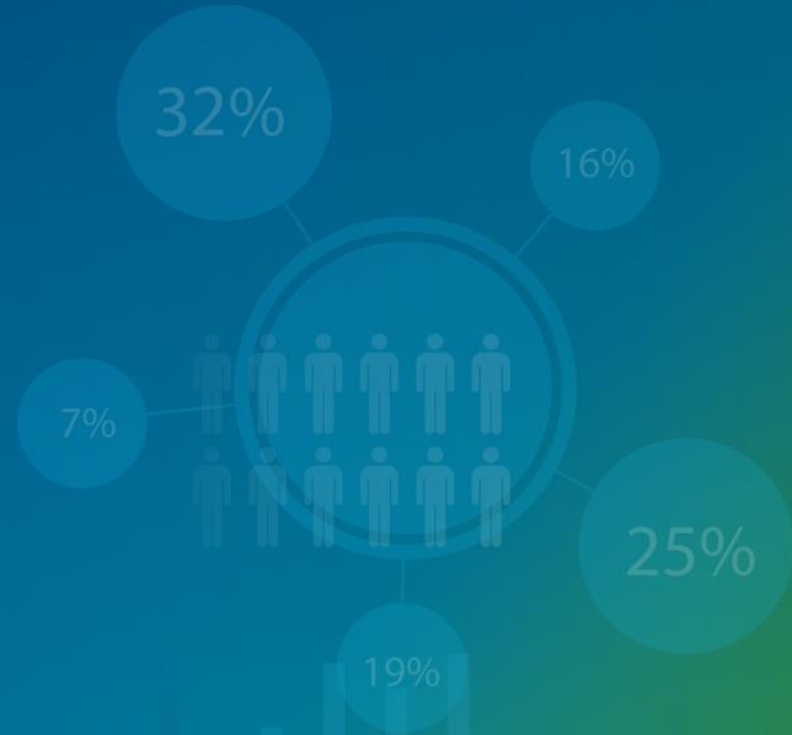
# Survey Experience Affects Participation Status (cont.)

## Association with time to withdraw after consent

- Survival tree



# Conclusions



# Conclusions

- Respondent burden
  - No particular questions triggered breakoff or changes of responses
  - Most of task components had reasonable length
- Contact attempt
  - Reaching out to participants helped increase completion rate
- Survey experience on participation status
  - Lower retention rate for index cases than that for household members
  - Effort on the previous components predicted response propensity of the next step
- More to explore with paradata

# Thank you

Xiaoshu Zhu

[xiaoshuzhu@westat.com](mailto:xiaoshuzhu@westat.com)

[westat.com](https://www.westat.com)

